



ORIGIN GGMDFCHFFRTPLPFAWRKQVQAFILCLITTAFLYLMRL"

Query Match 100.0%; Score 887; DB 9; Length 887;

Best Local Similarity 100.0%; Pred. No. 5.7e-144; Indels 0; Gaps 0;

Matches 887; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 CGGGCCAAAGAAACGACGAGAGCCCGGCCCCCGACAGAGGCGGAGCCATGGTTAGCA 60
DB 1 CGGGCCAAAGAAACGACGAGAGCCCGGCCCCCGACAGAGGCGGAGCCATGGTTAGCA 60
OY 61 GTTGGGGAGGCGACACCATGAGCCCGCTGCGGAGCGCACAGAGCTGTTGTGCG 120
DB 61 GTTGGGGAGGCGACACCATGAGCCCGCTGCGGAGCGCACAGAGCTGTTGTGCG 120
OY 121 CGACTACCTGGGGGTAAGTGGCCCGGAGACCCCGACCCCGAGCCGCGCATCCACGCC 180
DB 121 CGACTACCTGGGGGTAAGTGGCCCGGAGACCCCGACCCCGAGCCGCGCATCCACGCC 180
OY 181 CGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
OY 241 CTCGCGCTACCTGCGCTACCCCGGAGACCGCTTGCAGCTGCTGCTGCTGCTGCTGCT 300
DB 241 CTCGCGCTACCTGCGCTACCCCGGAGACCGCTTGCAGCTGCTGCTGCTGCTGCTGCT 300
OY 301 CGTGTCTCCGACAGCCCGGCGCCGACCTGGGGAGAGTGTGACCTGCTGACCTTGC 360
DB 301 CGTGTCTCCGACAGCCCGGCGCCGACCTGGGGAGAGTGTGACCTGCTGACCTTGC 360
OY 361 AGGAGCGCTGTGAGAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAAATGGGGCTTCCA 420
DB 361 AGGAGCGCTGTGAGAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAAATGGGGCTTCCA 420
OY 421 GCGCGGCTTAAAGAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAAATGGGGCTTCCA 480
DB 421 GCGCGGCTTAAAGAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAAATGGGGCTTCCA 480
OY 481 GGTGAGCTGCGGCTGATGAGAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAAATGGGG 540
DB 481 GGTGAGCTGCGGCTGATGAGAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAAATGGGG 540
OY 541 TGGCTTTTGTCACTTCTTCAAGAACCCCTTTCACAGCTGCTTTTGGAGAAAACAGTGT 600
DB 541 TGGCTTTTGTCACTTCTTCAAGAACCCCTTTCACAGCTGCTTTTGGAGAAAACAGTGT 600
OY 601 CAGGCTTTTGTCACTTCTTCAAGAACCCCTTTCACAGCTGCTTTTGGAGAAAACAGT 660
DB 601 CAGGCTTTTGTCACTTCTTCAAGAACCCCTTTCACAGCTGCTTTTGGAGAAAACAGT 660
OY 661 ATGAGTTTAAACCTTTTAAACCCGCTTTCACAGCTGCTTTCACAGCTTTCACAG 720
DB 661 ATGAGTTTAAACCTTTTAAACCCGCTTTCACAGCTGCTTTCACAGCTTTCACAG 720
OY 721 ATGTGTGAGAAACAAGAACTGAGGAAAGACCTTCCCAAGCTTTTATCTGA 780
DB 721 ATGTGTGAGAAACAAGAACTGAGGAAAGACCTTCCCAAGCTTTTATCTGA 780
OY 781 ATGCATACAGAGAGCTGAGAGGTGATTTGGCCAGTGTTTTAACTTGTACAGTACT 840
DB 781 ATGCATACAGAGAGCTGAGAGGTGATTTGGCCAGTGTTTTAACTTGTACAGTACT 840
OY 841 CAGGTGTGAGCAAGAAATGCAATGGCTTCTTCAAGTGAAGAA 887
DB 841 CAGGTGTGAGCAAGAAATGCAATGGCTTCTTCAAGTGAAGAA 887
```

RESULT 2  
LOCUS BD233466 1168 bp DNA linear PAT 17-JUL-2003  
DEFINITION Human protein having hydrophobic domain and DNA encoding the same.  
ACCESSION BD233466

VERSION BD233466.1 GI:33043236

KEYWORDS JP 2002519016-A/12.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE Kato, S. and Kimura, T. (bases 1 to 1168)

Human protein having hydrophobic domain and DNA encoding the same

JOURNAL Parent: JP 2002519016-A 12 02-JUL-2002;

SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC

OS Homo sapiens (human)

PN JP 2002519016-A/12

PD 02-JUL-2002

PF 18-JUN-1999 JP 200057267

PI SEISHI KATO, TOMOKO KIMURA

PC C12N15/09, C07K14/47, C12N1/15, C12N1/19, C12N5/10, C12N5/00, C12N5/PC

00

CC Human protein having hydrophobic domain and DNA encoding the

CC Key same Location/Qualifiers

FT source 1..1168

FT location/Qualifiers

FT 1..1168 /organism='Homo sapiens (human)'

FT /mol\_type='genomic DNA'

FT /db\_xref='taxon:9606'

ORIGIN

Query Match 91.6%; Score 812.4; DB 6; Length 1168;

Best Local Similarity 99.9%; Pred. No. 4.8e-131; Indels 1; Gaps 0;

Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY 74 ACCACCATGCGCGACCCGCTGCGGAGCGCACCGAGCTGTGCTGCGGACTACTGGGG 133
DB 1 ACCACCATGCGCGACCCGCTGCGGAGCGCACCGAGCTGTGCTGCGGACTACTGGGG 133
OY 134 TACTGCGCGCGGAGAACCGGAGACCCCGAGCGGAGCCATCCAGCCGAGGCGCGCG 193
DB 134 TACTGCGCGCGGAGAACCGGAGACCCCGAGCGGAGCCATCCAGCCGAGGCGCGCG 193
OY 61 TACTGCGCGCGGAGAACCGGAGACCCCGAGCGGAGCCATCCAGCCGAGGCGCGCG 120
DB 61 TACTGCGCGCGGAGAACCGGAGACCCCGAGCGGAGCCATCCAGCCGAGGCGCGCG 120
OY 194 CTGCGCTCGCGCGGCGGAGGTTACGAGATTCACCGGCTCTTTTCTCGGCTACCTC 253
DB 194 CTGCGCTCGCGCGGCGGAGGTTACGAGATTCACCGGCTCTTTTCTCGGCTACCTC 253
OY 121 CTGCGCTCGCGCGGCGGAGGTTACGAGATTCACCGGCTCTTTTCTCGGCTACCTC 180
DB 121 CTGCGCTCGCGCGGCGGAGGTTACGAGATTCACCGGCTCTTTTCTCGGCTACCTC 180
OY 254 GGTACCCCGGAGAACCGCTTTCAGAGCTGTGCGCTGATGCGGATTCGCTTCCGAC 313
DB 254 GGTACCCCGGAGAACCGCTTTCAGAGCTGTGCGCTGATGCGGATTCGCTTCCGAC 313
OY 181 GGTACCCCGGAGAACCGCTTTCAGAGCTGTGCGCTGATGCGGATTCGCTTCCGAC 240
DB 181 GGTACCCCGGAGAACCGCTTTCAGAGCTGTGCGCTGATGCGGATTCGCTTCCGAC 240
OY 314 AGCCCGGCGGAGAACCGCTTTCAGAGCTGTGCGCTGATGCGGATTCGCTTCCGAC 373
DB 314 AGCCCGGCGGAGAACCGCTTTCAGAGCTGTGCGCTGATGCGGATTCGCTTCCGAC 373
OY 241 AGCCCGGCGGAGAACCGCTTTCAGAGCTGTGCGCTGATGCGGATTCGCTTCCGAC 300
DB 241 AGCCCGGCGGAGAACCGCTTTCAGAGCTGTGCGCTGATGCGGATTCGCTTCCGAC 300
OY 374 GAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAGAGTGGGCTTCCAGCCGCGCTAAG 433
DB 374 GAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAGAGTGGGCTTCCAGCCGCGCTAAG 433
OY 301 GAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAGAGTGGGCTTCCAGCCGCGCTAAG 360
DB 301 GAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAGAGTGGGCTTCCAGCCGCGCTAAG 360
OY 434 GAGAGAGGCGCGAGTGTGCGCGCGGAGCTGCGAGCGCTGTGAGCTTGTGAC 493
DB 434 GAGAGAGGCGCGAGTGTGCGCGCGGAGCTGCGAGCGCTGTGAGCTTGTGAC 493
OY 361 GAGAGAGGCGCGAGTGTGCGCGCGGAGCTGCGAGCGCTGTGAGCTTGTGAC 420
DB 361 GAGAGAGGCGCGAGTGTGCGCGCGGAGCTGCGAGCGCTGTGAGCTTGTGAC 420
OY 494 CTCATGAGGAGAGACCGGCGGCTGCGAGGCTGCGAGGCGGCTGGGATGGCTTGTGAC 553
DB 494 CTCATGAGGAGAGACCGGCGGCTGCGAGGCTGCGAGGCGGCTGGGATGGCTTGTGAC 553
OY 421 CTCATGAGGAGAGACCGGCGGCTGCGAGGCTGCGAGGCGGCTGGGATGGCTTGTGAC 480
DB 421 CTCATGAGGAGAGACCGGCGGCTGCGAGGCTGCGAGGCGGCTGGGATGGCTTGTGAC 480
OY 554 TTCTTCAAGAGACCCCTTTCACAGCTGTGTTTGGAGAAAACAGCTGTCCAGGCTTTCTG 613
DB 554 TTCTTCAAGAGACCCCTTTCACAGCTGTGTTTGGAGAAAACAGCTGTGTCCAGGCTTTCTG 613
OY 481 TTCTTCAAGAGACCCCTTTCACAGCTGTGTTTGGAGAAAACAGCTGTGTCCAGGCTTTCTG 540
DB 481 TTCTTCAAGAGACCCCTTTCACAGCTGTGTTTGGAGAAAACAGCTGTGTCCAGGCTTTCTG 540
OY 614 TCATGCTGTGTTAACAACAGCTTCAATTAATCTGAGACAGATTAATGAAGTTTAA 673
DB 614 TCATGCTGTGTTAACAACAGCTTCAATTAATCTGAGACAGATTAATGAAGTTTAA 673
```

Db 541 TCATGCTTTTAACAACAGCCCTTCAATTATCTGTGAGACAGATTATTAAGTTTTAAAA 600  
Qy 674 CTTTAAACCCGCTTACCTGCGCAACTGTGACCAACTAAATGACAGATGTGTGAGACA 733  
Db 601 CTTTAAACCCGCTTACCTGCGCAACTGTGACCAACTAAATGACAGATGTGTGAGACA 660-715  
Qy 734 AGAAGTGAAGGAAAGACCTTCCCGACCCGACGAGAGCTTTTATCTGAATGACATACAGAA 793  
Db 661 AGAAGTGAAGGAAAGACCTTCCCGACCCGACGAGAGCTTTTATCTGAATGACATACAGAA 720  
Qy 794 GTCTGAGGTGTGTATTTGGCCAGTGTTTTAACTTTGTGACAGTACTAGGTGTGAGAC 853  
Db 721 GTCTGAGGTGTGTATTTGGCCAGTGTTTTAACTTTGTGACAGTACTAGGTGTGAGAC 780  
Qy 854 AAGATGCAATGCTCTCTTGTGAGTGAAGAA 887  
Db 781 AAGATGCAATGCTCTCTTGTGAGTGAAGAA 814

RESULT 3  
AF285092 615 bp mRNA linear PRI 08-NOV-2001  
LOCUS AF285092 Homo sapiens Bcl-2-like protein 10 mRNA, complete cds.  
DEFINITION AF285092 AF285092.1 GI:9837265  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
Zhang, H., Holzgreve, W. and De Geyter, C.  
Bcl2-L-10, a novel anti-apoptotic member of the Bcl-2 family, blocks apoptosis in the mitochondria death pathway but not in the death receptor pathway  
Hum. Mol. Genet. 10 (21), 2329-2339 (2001)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
Zhang, H.H.  
Direct Submission  
Submitted (05-JUL-2000) University Women's Hospital, Schanzenstrasse 46, Basel 4057, Switzerland  
JOURNAL  
FEATURES  
source  
location/Qualifiers  
1. 615  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q21"  
/tissue\_type="ovary"  
1. 615  
/codon\_start=1  
/product="Bcl-2-like protein 10"  
/protein\_id="AA00503.1"  
/db\_xref="GI:9837265"  
/translation="MWDLRRTTMDPLRRTLLADYGYCARERTPEPASTP  
EAAVLRSAARLQRIHRSFSAVIGYENRRLVADSLSDSPGRTWRVTVLT  
PAGTLRGLPLVTRMKWQFQPRLEKQGVARDCCQLVALSLSRLLSGHRAWLOAQ  
GMDGFCFFFTFPPLAFWRKQLOVAFISCLITTAFTLMTLL"

CDS  
1. 615  
/product="Bcl-2-like protein 10"  
/protein\_id="AA00503.1"  
/db\_xref="GI:9837265"  
/translation="MWDLRRTTMDPLRRTLLADYGYCARERTPEPASTP  
EAAVLRSAARLQRIHRSFSAVIGYENRRLVADSLSDSPGRTWRVTVLT  
PAGTLRGLPLVTRMKWQFQPRLEKQGVARDCCQLVALSLSRLLSGHRAWLOAQ  
GMDGFCFFFTFPPLAFWRKQLOVAFISCLITTAFTLMTLL"

ORIGIN  
Query Match 69.3%; Score 615; DB 9; Length 615;  
Best Local Similarity 100.0%; Pred. No. 8.3e-97;  
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 ATGGTTGACGAGTTCGGGAGGCGACACATGCGGACCGCTCGGGAGGCGACCCAG 109  
Db 1 ATGGTTGACGAGTTCGGGAGGCGACACATGCGGACCGCTCGGGAGGCGACCCAG 60  
Qy 110 CTGTTGCTGGCGAGTACTCTGGGAGTCTGCGCCCGGAAACCGCGACCCCGAGCGGCG 169  
Db 61 CTGTTGCTGGCGAGTACTCTGGGAGTCTGCGCCCGGAAACCGCGACCCCGAGCGGCG 120

Qy 170 CCATCAGCGCCGAGGCGCGGCTGCTGCGCGCGCGCGAGGTTACGCGAGATTTCAC 229  
Db 121 CCATCAGCGCCGAGGCGCGGCTGCTGCGCGCGCGCGAGGTTACGCGAGATTTCAC 180  
Qy 230 CGGTCTTTTTCCTCGCTTACCTTGGCTAACCCCGGAAACGCTTTCAGCTGTGGCGCTG 289  
Db 181 CGGTCTTTTTCCTCGCTTACCTTGGCTAACCCCGGAAACGCTTTCAGCTGTGGCGCTG 240  
Qy 290 ATGGCGGATTCGCGCTCTCCGACAGCCCGCGCCACCTGGGGCAGAGTGTGACGCTC 349  
Db 241 ATGGCGGATTCGCGCTCTCCGACAGCCCGCGCCACCTGGGGCAGAGTGTGACGCTC 300  
Qy 350 GTGACCTTCGCGAGGACGCTCTGAGAGAGGCGCGCTGTGACCGCGCGGTGAGAAAG 409  
Db 301 GTGACCTTCGCGAGGACGCTCTGAGAGAGGCGCGCTGTGACCGCGCGGTGAGAAAG 360  
Qy 410 TGGGCGCTTCGAGCGCGGCTTAAAGAGAGGAGGAGCGTCCCGGAGCTGCCAGCGC 469  
Db 361 TGGGCGCTTCGAGCGCGGCTTAAAGAGAGGAGGAGCGTCCCGGAGCTGCCAGCGC 420  
Qy 470 CTGGTGGCTTGTGAGCTCGCGGCTCATGTGGGGCAGACCGCGCTGTGAGAGCTCAG 529  
Db 421 CTGGTGGCTTGTGAGCTCGCGGCTCATGTGGGGCAGACCGCGCTGTGAGAGCTCAG 480  
Qy 530 GCGCGCTGGAGTGGCTTTTGTCACTTCTTCAGAGCCCGCTTCACCTGCTTTTGGAGA 589  
Db 481 GCGCGCTGGAGTGGCTTTTGTCACTTCTTCAGAGCCCGCTTCACCTGCTTTTGGAGA 540  
Qy 590 AAACAGCTGTTCAGAGCTTTTCTGTCACTTGTAAACAGAGCTTATTTATCTCGG 649  
Db 541 AAACAGCTGTTCAGAGCTTTTCTGTCACTTGTAAACAGAGCTTATTTATCTCGG 600  
Qy 650 ACACAGTTATATGA 664  
Db 601 ACACAGTTATATGA 615

RESULT 4  
HSA458330 585 bp DNA linear PRI 25-APR-2002  
LOCUS HSA458330 Homo sapiens NRH gene for anti-apoptotic protein.  
DEFINITION AJ458330 AJ458330.1 GI:20338765  
ACCESSION  
VERSION  
KEYWORDS  
anti-apoptotic protein, NRH gene.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
1  
Aouacheria, A., Arnaud, R., Venet, S., Lalle, P., Gouy, M., Rigal, D. and Gillet, G.  
Nrh, a human homologue of Nr-13 associates with Bcl-Xs and is an inhibitor of apoptosis  
Oncogene 20 (41), 5846-5855 (2001)  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
Gillet, G.  
Direct Submission  
Submitted (23-APR-2002) Gillet G., Ibcip, CNRS UMR 5086, 7 PASSAGE DU VERCORS, 69367, FRANCE  
JOURNAL  
FEATURES  
source  
location/Qualifiers  
1. 585  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
1. 585  
/gene="NRH"  
1. 585  
/function="apoptosis inhibitor"

```
/codon_start=1
/product="anti-apoptotic protein"
/protein_id="CAD30221.1"
/db_xref="gi:2038766"
/db_xref="GOL:09HD36"
/db_xref="UniProt/Swiss-Prot:Q9HD36"
/translation="MADPLRRETELMLADYLCARBPETEPAPSTPEAAVLRSAA
RLROHRSFSAVLGYPNRPFELVALMADSVLSDPGFTWGVTLVTFAGTILERP
LVTARMKGFOPRLKEQGVARDQRLVALLSRLMGQRHAWIQAGGMDGFCHF
RTPPLAWRQQLVQAFISCLITTAFTYLMTRLL"
```

## ORIGIN

```
Query Match      66.0%; Score 585; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.3e-91;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      80 ATGGCCGACCCGCTCGGGAGGCGACCGAGCTGTTGCTGCGGACTACTTGGGGTACTGC 139
Db       1 ATGGCCGACCCGCTCGGGAGGCGACCGAGCTGTTGCTGCGGACTACTTGGGGTACTGC 60

QY      140 GCCCGGAAACCCGCGACCCCGAGCGCGCATCCAGCCCGAGGCGCGTGTGCGGC 199
Db       61 GCCCGGAAACCCGCGACCCCGAGCGCGCATCCAGCCCGAGGCGCGTGTGCGGC 120

QY      200 TCCGCGCGCGCAGGTTACGCGAGATTACCGGTCCTTTTCTCCGCTTACCTCGGCTAC 259
Db       121 TCCGCGCGCGCAGGTTACGCGAGATTACCGGTCCTTTTCTCCGCTTACCTCGGCTAC 180

QY      260 CCGCGGAAACCGCTTGGAGTGTGTGCGCTGATGTGGGATTCGTGTCTTCCGACAGCCGC 319
Db       181 CCGCGGAAACCGCTTGGAGTGTGTGCGCTGATGTGGGATTCGTGTCTTCCGACAGCCGC 240

QY      320 GGGCCCACTGGGGGAGAGTGTGACGCTGTCGACCTTCGACGAGGAGCGTGTGAGAGA 379
Db       241 GGGCCCACTGGGGGAGAGTGTGACGCTGTCGACCTTCGACGAGGAGCGTGTGAGAGA 300

QY      380 GGGCCGCTGTGTACCGCCCGGTGAGAGAGTGGGGCTTTCAGCCCGGCTAAAGAGCAG 439
Db       301 GGGCCGCTGTGTACCGCCCGGTGAGAGAGTGGGGCTTTCAGCCCGGCTAAAGAGCAG 360

QY      440 GAGGGCGACGTGTGCGCCCGGAGCTGCTGTGTGCTTGTGCTGCTGCTGCTGCTGCTG 499
Db       361 GAGGGCGACGTGTGCGCCCGGAGCTGCTGTGTGCTTGTGCTGCTGCTGCTGCTGCTG 420

QY      500 GGGCGACACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 559
Db       421 GGGCGACACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY      560 AGGACCCCTTTCACCTGCTTTCGAGAGAAACAGCTGTGTCAGGCTTTCGTGCATGC 619
Db       481 AGGACCCCTTTCACCTGCTTTCGAGAGAAACAGCTGTGTCAGGCTTTCGTGCATGC 540

QY      620 TTGTTAACACAGCCTTCACTTATCTCTGAGACAGATTATATGA 664
Db       541 TTGTTAACACAGCCTTCACTTATCTCTGAGACAGATTATATGA 585
```

RESULT 5  
BD233456 582 bp DNA linear PAT 17-JUL-2003

DEFINITION Human protein having hydrophobic domain and DNA encoding the same.

ACCESSION BD233456

VERSION BD233456.1 GI:33043226

KEYWORDS JP 2002519016-A/2.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 582)

Kato, S. and Kimura, T.

Human protein having hydrophobic domain and DNA encoding the same

Patient: JP 2002519016-A 2 02-JUL-2002.

JOURNAL SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC

## COMMENT

OS Homo sapiens (human)

PN JP 2002519016-A/2

PD 02-JUL-2002

PP 18-JUN-1999 JP 2000557267

PI SEISHI KATO, TOMOKO KIMURA

PC C12N15/09, C07K14/47, C12N1/15, C12N1/19, C12N5/10, C12N5/00, C12N5/00

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the

CC Human protein having hydrophobic domain and DNA encoding the



ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kites, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 38039 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
Source Location/Qualifiers  
1. .726  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

Query Match 64.8%; Score 574.4; DB 6; Length 726;  
Best Local Similarity 90.2%; Pred. No. 9e-90;  
Matches 655; Conservative 0; Mismatches 1; Indels 70; Gaps 1;

QY 50 ATGGTTGACCAAGTTCGGGAGGAGCCACCATGAGCCGCTGCGGAGGAGCCAG 109  
DB 1 ATGGTTGACCAAGTTCGGGAGGAGCCACCATGAGCCGCTGCGGAGGAGCCAG 60  
QY 110 CTGTTGCTGGGCGACTACCTGAGGAGTCTGCGGAGGAGCCGAGCCGAGCCGCG 169  
DB 61 CGGTTGCTGGGCGACTACCTGAGGAGTCTGCGGAGGAGCCGAGCCGAGCCGCG 120  
QY 170 CCATCCAGCCGAGGAGCCGAGTCTGCGGAGGAGCCGAGGAGTTCAGCAATTAC 229  
DB 121 CCATCCAGCCGAGGAGCCGAGTCTGCGGAGGAGCCGAGGAGTTCAGCAATTAC 180  
QY 230 CGGTTCTTTTCTGCGCTACCTGAGGAGTCTGCGGAGGAGCCGTTGAGCTGCGCTG 289  
DB 181 CGGTTCTTTTCTGCGCTACCTGAGGAGTCTGCGGAGGAGCCGTTGAGCTGCGCTG 240  
QY 290 ATGGCGGAGTTCGAGTCTGCGAGAGCCGAGCCGAGCCGAGGAGGAGTTCAGCTC 349  
DB 241 ATGGCGGAGTTCGAGTCTGCGAGAGCCGAGCCGAGCCGAGGAGGAGTTCAGCTC 300  
QY 350 GTGACCTTCGAGGAGGAGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 409  
DB 301 GTGACCTTCGAGGAGGAGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
QY 410 TGGGAGCTTCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 469  
DB 361 TGGGAGCTTCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
QY 470 CTGTTGCTGGGCGACTACCTGAGGAGTCTGCGGAGGAGCCGAGGAGGAGGAGGAG 529  
DB 421 CTGTTGCTGGGCGACTACCTGAGGAGTCTGCGGAGGAGCCGAGGAGGAGGAGGAG 480  
QY 530 GAGCGAGCTG----- 537  
DB 481 GAGCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
QY 538 -----GAGTGGCTTTTTCAGCTTTCAGAGAGCCCTTTTTCAGCTGGC 579  
DB 541 CGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
QY 580 TTTTGGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 639  
DB 601 TTTTGGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
QY 640 TTATCTCTGAGCAGATTTATGAGTTTAAACTTTTAAACCCGCTTCTACCTGCCAA 699  
DB 661 TTATCTCTGAGCAGATTTATGAGTTTAAACTTTTAAACCCGCTTCTACCTGCCAA 720  
QY 700 CTGTGA 705  
DB 721 CTGTGA 726

RESULT 7  
AC023906/c  
LOCUS  
DEFINITION Homo sapiens chromosome 15 clone CTD-2184D3 map 15q21.2, complete  
sequence.  
ACCESSION AC023906  
VERSION AC023906.7  
KEYWORDS HTG.  
SOURCE  
ORGANISM Homo sapiens (human)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AC023906 93287 bp DNA linear PRI 06-JUL-2001  
Homo sapiens chromosome 15 clone CTD-2184D3 map 15q21.2, complete  
sequence.  
AC023906.7 GI:14595770  
HTG.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 93287)  
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,  
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,  
Pate, D., and Hood, L.  
Sequencing of human chromosome 15 D15S146-D15S117 region  
Unpublished  
2 (bases 1 to 93287)  
Rowen, L., Madan, A., Qin, S., Abbaei, N., Baradarani, L., Birditt, B.,  
Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G.,  
James, R., Kaur, A., Madan, A., Owen, M. P., Ratcliffe, A., Shaffer, T.,  
and Hood, L.  
Direct Submission  
Submitted (20-FEB-2000) Multimegabase Sequencing Center, University  
of Washington, PO BOX 357730, Seattle, WA 98195, USA  
3 (bases 1 to 93287)  
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,  
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,  
Pate, D., and Hood, L.  
Direct Submission  
Submitted (04-JUL-2001) Multimegabase Sequencing Center, Institute  
for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA  
98105, USA  
4 (bases 1 to 93287)  
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,  
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,  
Pate, D., and Hood, L.  
Direct Submission  
Submitted (06-JUL-2001) Multimegabase Sequencing Center, Institute  
for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA  
98105, USA  
On Jul 4, 2001 this sequence version replaced gi:12248292.

Genome Center  
Center: Multimegabase Sequencing Center  
Center code: UMWSC  
Web site: [http://chroma.mbc.washington.edu/msg\\_www](http://chroma.mbc.washington.edu/msg_www)  
Contact: [leerowen@systemsbiology.org](mailto:leerowen@systemsbiology.org)

Summary Statistics  
Sequencing vector: pUC18; L08752  
Chemistry: Dye-terminator Big Dye; 90% of reads  
Chemistry: Dye-primer Big Dye; 10% of reads  
Assembly program: Phrap; version 0.990399

Note: Data from overlapping BACs AC010674 [drafting center:  
UMMSC], AC090970 [drafting center: UMWSC], and AC016824 [drafting  
center: GTC] was added for finishing

FEATURES  
Source  
1. .93287  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q21.2"  
/clone="CTD-2184D3"  
/clone\_lib="Cal Tech Human BAC library D"  
/note="Data from overlapping clones CTD-2650P22 AC090970,  
RP11-47X1 AC016824, and RP11-430B1 was added and the  
consensus was determined from CTD-2184D3 to the extent  
possible"  
1. .8893  
misc\_feature

	Query Match	60.8%; Score 539.2; DB 9; Length 93287;
	Best Local Similarity	99.4%; Pred. No. 7,7e-84;
	Matches 541; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
misc_feature	/note="overlap with CTD-2650P22 AC09070"	
unure	1. .8893	
unure	/note="overlap with RP11-47K1 AC016824"	
unure	6078. .6140	
unure	/note="low quality data"	
unure	10529. .10532	
unure	/note="low quality data"	
unure	14722. .14808	
unure	/note="low quality data"	
unure	37627	
unure	/note="low quality data"	
unure	71552. .71593	
unure	/note="low quality data"	
unure	82012. .93287	
misc_feature	/note="overlap with RP11-430B1 AC010674"	
ORIGIN		
QY	1	CGGGCCAAAGAAACACAGGAGGCGCGGCCCGCCAGCAGAGGCGGAGCCATGGTTTACCA 60
DB	64991	CGGGCCAAAGAAACACAGGAGGCGCGGCCCGCCAGCAGAGGCGGAGCCATGGTTTACCA 649322
QY	61	GTTGCGGAGGCGCACACATGCGCAACCGCTGCGGAGCGCACCGAGCTGTTGCTGAC 120
DB	64931	GTTGCGGAGGCGCACACATGCGCAACCGCTGCGGAGCGCACCGAGCTGTTGCTGAC 648722
QY	121	CGACTACCTGGGGATACCTGCGCGCGGAAACCGCGACCCCGAGCCCGGCGCATTCACGCG 180
DB	64871	CGACTACCTGGGGATACCTGCGCGCGGAAACCGCGACCCCGAGCCCGGCGCATTCACGCG 648122
QY	181	CGAGGCGCGCTGCTGCGCTCCGCGGCGCAGGTTACGGCAGATTACCGGATCTTTT 240
DB	64811	CGAGGCGCGCTGCTGCGCTCCGCGGCGCAGGTTACGGCAGATTACCGGATCTTTT 647522
QY	241	CTCGGCTTACTCGGCTTACCCCGGGAACCGCTTGAAGTGTGTGCGTGTGCGGATTC 300
DB	64751	CTCGGCTTACTCGGCTTACCCCGGGAACCGCTTGAAGTGTGTGCGTGTGCGGATTC 646922
QY	301	CGTGCCTCCGACAGCCCGGCGCCCACTGAGGGGAGAGTGTGACGCTGTAACCTTCCG 360
DB	64691	CGTGCCTCCGACAGCCCGGCGCCCACTGAGGGGAGAGTGTGACGCTGTAACCTTCCG 646322
QY	361	AGGACCGCTGAGAGAGAGGCGCGTGTGACCGCCCGGTGAGAAAGTGTGGGCTTCCA 420
DB	64631	AGGACCGCTGAGAGAGAGGCGCGTGTGACCGCCCGGTGAGAAAGTGTGGGCTTCCA 645722
QY	421	GCCCGGCTTAAAGAGAGAGAGGCGCGACGTGCGCCCGGGAATTGCAAGCGCTGTGGCCCTT 480
DB	64571	GCCCGGCTTAAAGAGAGAGAGGCGCGACGTGCGCCCGGGAATTGCAAGCGCTGTGGCCCTT 645122
QY	481	GCTGAGCTCGCGGCTTACGAGGCGACCGCGCTGAGGCTTCAAGGCGGCTGGGAG 540
DB	64511	GCTGAGCTCGCGGCTTACGAGGCGACCGCGCTGAGGCTTCAAGGCGGCTGGGAG 644522
QY	541	TGGC 544
DB	64451	GAGC 64448
RESULT 8		
LOCUS	AC018903	214669 bp DNA 1linear HTG 04-JUN-2000
DEFINITION	Homo sapiens chromosome 15 clone RP11-337B11 map 15q21, LOW-PASS	
ACCESSION	AC018903	SEQUENCE SAMPLING.
VERSION	AC018903.2	GI:8247797
KEYWORDS	HTG; HTGS; PHASPO.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

```

REFERENCE Mammalia: Eutheria: Primates, Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 214669)
Rowen, L., Madan, A., Qin, S., Abbasi, N., Baradarani, L., Birditt, B.,
Bloom, S., Dots, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A.,
Madan, A., Nesbitt, R., Shaffer, T. and Hood, L.
JOURNAL Sequencing of human chromosome 15 D15S146-D15S147 region
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 214669)
Rowen, L., Madan, A., Qin, S., Abbasi, N., Baradarani, L., Birditt, B.,
Bloom, S., Dots, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A.,
James, R., Kaur, A., Madan, A., Owen, M.P., Ratcliffe, A., Shaffer, T.
and Hood, L.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-1999) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT On Jun 4, 2000 this sequence version replaced gi:6630517.
-----Genome Center-----
Center: Multimegabase Sequencing Center
Center code: UMWSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leerowen@systembiology.org
-----Summary Statistics-----
Sequencing vector: pUC18, 108752
Chemistry: Dye-terminator Big Dye, 90% of reads
Chemistry: Dye-primer Big Dye, 10% of reads
Assembly program: Phrap, version 0.990399
-----
* NOTE: This record contains 192 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1
* 806: contig of 806 bp in length
* 807 906: gap of unknown length
* 907 2068: contig of 1162 bp in length
* 2069 2168: gap of unknown length
* 2169 3020: contig of 852 bp in length
* 3021 3120: gap of unknown length
* 3121 4284: contig of 1164 bp in length
* 4285 4384: gap of unknown length
* 4385 5201: contig of 817 bp in length
* 5202 5301: gap of unknown length
* 5302 6544: contig of 1243 bp in length
* 6545 6645: gap of unknown length
* 6645 7472: contig of 828 bp in length
* 7473 7572: gap of unknown length
* 7573 8756: contig of 1184 bp in length
* 8757 8856: gap of unknown length
* 8857 9658: contig of 802 bp in length
* 9659 9759: gap of unknown length
* 9759 10905: contig of 1147 bp in length
* 10906 11005: gap of unknown length
* 11006 11811: contig of 806 bp in length
* 11812 11911: gap of unknown length
* 11912 13293: contig of 1382 bp in length
* 13294 13393: gap of unknown length
* 13394 14211: contig of 818 bp in length
* 14212 14311: gap of unknown length
* 14312 15465: contig of 1154 bp in length
* 15466 15565: gap of unknown length
* 15566 16379: contig of 814 bp in length
* 16380 16479: gap of unknown length
* 16480 17624: contig of 1145 bp in length
* 17625 17724: gap of unknown length
* 17725 18578: contig of 854 bp in length
* 18579 18678: gap of unknown length
* 18679 19774: contig of 1096 bp in length

```

*	19775	19874: gap of unknown length	*	59846	60655: contig of 810 bp in length
*	19875	20700: contig of 826 bp in length	*	60755	60755: gap of unknown length
*	20701	20800: gap of unknown length	*	60756	61976: contig of 1221 bp in length
*	20801	22036: contig of 1236 bp in length	*	61977	62076: gap of unknown length
*	22037	22136: gap of unknown length	*	62077	62896: contig of 820 bp in length
*	22137	23023: contig of 887 bp in length	*	62897	62986: gap of unknown length
*	23024	23123: gap of unknown length	*	62997	64128: contig of 1132 bp in length
*	23124	24196: contig of 1073 bp in length	*	64129	64228: gap of unknown length
*	24197	24296: gap of unknown length	*	64229	65026: contig of 798 bp in length
*	24297	25141: contig of 845 bp in length	*	65027	65126: gap of unknown length
*	25142	25241: gap of unknown length	*	65127	66269: contig of 1143 bp in length
*	25242	26385: contig of 1144 bp in length	*	66270	66369: gap of unknown length
*	26386	26485: gap of unknown length	*	66370	67220: contig of 851 bp in length
*	26486	27325: contig of 840 bp in length	*	67221	67320: gap of unknown length
*	27326	27425: gap of unknown length	*	67321	68425: contig of 1105 bp in length
*	27426	28691: contig of 1266 bp in length	*	68426	68525: gap of unknown length
*	28692	28791: gap of unknown length	*	68526	69322: contig of 797 bp in length
*	28792	29655: contig of 864 bp in length	*	69323	69422: gap of unknown length
*	29656	29755: gap of unknown length	*	69423	70568: contig of 1146 bp in length
*	29756	30681: contig of 1126 bp in length	*	70569	70668: gap of unknown length
*	30682	30981: gap of unknown length	*	70669	71547: contig of 879 bp in length
*	30982	31813: contig of 832 bp in length	*	71548	71647: gap of unknown length
*	31814	31913: gap of unknown length	*	71648	73629: contig of 1982 bp in length
*	31914	33056: contig of 1143 bp in length	*	73630	73729: gap of unknown length
*	33057	33156: gap of unknown length	*	73730	74587: contig of 858 bp in length
*	33157	34007: contig of 851 bp in length	*	74588	74687: gap of unknown length
*	34008	34107: gap of unknown length	*	74688	75780: contig of 1093 bp in length
*	34108	35320: contig of 1213 bp in length	*	75781	75880: gap of unknown length
*	35321	35420: gap of unknown length	*	75881	77196: contig of 1316 bp in length
*	35421	36207: contig of 787 bp in length	*	77197	77296: gap of unknown length
*	36208	36307: gap of unknown length	*	77297	78651: contig of 1355 bp in length
*	36308	37521: contig of 1214 bp in length	*	78652	78751: gap of unknown length
*	37522	37621: gap of unknown length	*	78752	79557: contig of 806 bp in length
*	37622	38420: contig of 799 bp in length	*	79558	79657: gap of unknown length
*	38421	38520: gap of unknown length	*	79658	80774: contig of 1117 bp in length
*	38521	39726: contig of 1206 bp in length	*	80775	80874: gap of unknown length
*	39727	39826: gap of unknown length	*	80875	81690: contig of 816 bp in length
*	39827	40635: contig of 809 bp in length	*	81691	81790: gap of unknown length
*	40636	40735: gap of unknown length	*	81791	82928: contig of 1138 bp in length
*	40736	41845: contig of 1110 bp in length	*	82929	83028: gap of unknown length
*	41846	41945: gap of unknown length	*	83029	83902: contig of 874 bp in length
*	41946	42730: contig of 785 bp in length	*	83903	84002: gap of unknown length
*	42731	42830: gap of unknown length	*	84003	85193: contig of 1191 bp in length
*	42831	43917: contig of 1087 bp in length	*	85194	85293: gap of unknown length
*	43918	44017: gap of unknown length	*	85294	86149: contig of 856 bp in length
*	44018	44825: contig of 808 bp in length			
*	44826	44925: gap of unknown length			
*	44926	46114: contig of 1189 bp in length			
*	46115	46214: gap of unknown length			
*	46215	47031: contig of 817 bp in length			
*	47032	47131: gap of unknown length			
*	47132	48292: contig of 1161 bp in length			
*	48293	48392: gap of unknown length			
*	48393	49176: contig of 784 bp in length			
*	49177	49276: gap of unknown length			
*	49277	50389: contig of 1113 bp in length			
*	50389	50489: gap of unknown length			
*	50490	51333: contig of 844 bp in length			
*	51334	51433: gap of unknown length			
*	51434	52765: contig of 1332 bp in length			
*	52766	52865: gap of unknown length			
*	52866	53710: contig of 845 bp in length			
*	53711	53810: gap of unknown length			
*	53811	54965: contig of 1155 bp in length			
*	54966	55065: gap of unknown length			
*	55066	55910: contig of 845 bp in length			
*	55911	56010: gap of unknown length			
*	56011	57558: contig of 1548 bp in length			
*	57559	57658: gap of unknown length			
*	57659	58509: contig of 851 bp in length			
*	58510	58609: gap of unknown length			
*	58610	59745: contig of 1136 bp in length			
*	59746	59845: gap of unknown length			

Query Match

Best Local Similarity 98.9%; Pred. No. 2.6e-50;

Matches 348; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	536	TGGAGTGGCTTTTGTCACTTCTTGAAGACCCCTTTTCACTGCTTTTGGAGAAACAG	595
DB	168582	TAGATGCTTTTGTCACTTCTTGAAGACCCCTTTTCACTGCTTTTGGAGAAACAG	168523
QY	596	CTGGTCCAGGCTTTTGTCACTTCTTGAAGACCCCTTTTCACTGCTTTTGGAGAAACAG	655
DB	168522	CTGGTCCAGGCTTTTGTCACTTCTTGAAGACCCCTTTTCACTGCTTTTGGAGAAACAG	168463
QY	656	TTATTATGAGTTTAAACCTTTTAAACCGCTTTTCACTGCTTTTGGAGAAACAG	715
DB	168462	TTATTATGAGTTTAAACCTTTTAAACCGCTTTTCACTGCTTTTGGAGAAACAG	168403
QY	716	GACAGATGTGTGAAGACAGACTGAGGAAAGACCTTCCCAACCCCAAGCGTTT	775
DB	168402	GACAGATGTGTGAAGACAGACTGAGGAAAGACCTTCCCAACCCCAAGCGTTT	168343
QY	776	TCTGAATGCAATCAAGAGGCTCTGAGGAGGATTTTGGCCAGTGTTTAACTTGAACA	835
DB	168342	TCTGAATGCAATCAAGAGGCTCTGAGGAGGATTTTGGCCAGTGTTTAACTTGAACA	168283
QY	836	GTACTCAGGTGTGAAGACAGAAATGCAATGCTCTTCTTGAAGTGAAGAA	887
DB	168282	GTACTCAGGTGTGAAGACAGAAATGCAATGCTCTTCTTGAAGTGAAGAA	168231

```
RESULT 9
AC018903
LOCUS
DEFINITION Homo sapiens chromosome 15 clone RP11-337B11 map 15q21, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION AC018903
VERSION AC018903.2 GI:8247797
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A.,
Madan, A., Nesbitt, R., Shaffer, T. and Hood, L.
1 (bases 1 to 214669)
2 (bases 1 to 214669)
TITLE
JOURNAL Unpublished
REFERENCE
AUTHORS Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A.,
Madan, A., Nesbitt, R., Shaffer, T. and Hood, L.
1 (bases 1 to 214669)
2 (bases 1 to 214669)
TITLE
JOURNAL Direct Submission
REFERENCE
AUTHORS Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A.,
Madan, A., Nesbitt, R., Shaffer, T. and Hood, L.
1 (bases 1 to 214669)
2 (bases 1 to 214669)
TITLE
JOURNAL Submitted (22-DEC-1999) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
On Jun 4, 2000 this sequence version replaced gi:6630517.
----- Genome Center
Center: Multimegabase Sequencing Center
Center code: UMMS
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leetowen@systemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-Primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
-----
* NOTE: This record contains 192 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 806: contig of 806 bp in length
* 807 906: gap of unknown length
* 907 2068: contig of 1162 bp in length
* 2069 2168: gap of unknown length
* 2169 3020: contig of 852 bp in length
* 3021 3120: gap of unknown length
* 3121 4284: contig of 1164 bp in length
* 4285 4384: gap of unknown length
* 4385 5201: contig of 817 bp in length
* 5202 5302: gap of unknown length
* 5302 6544: contig of 1243 bp in length
* 6545 7472: gap of unknown length
* 7473 7572: contig of 828 bp in length
* 7573 8756: gap of unknown length
* 8757 8856: gap of unknown length
* 8857 9658: gap of 802 bp in length
* 9659 9758: gap of unknown length
* 9759 10905: contig of 1147 bp in length
* 10906 11005: gap of unknown length
* 11006 11811: contig of 806 bp in length
* 11812 11911: gap of unknown length
*
* 11912 13293: contig of 1382 bp in length
* 13294 13393: gap of unknown length
* 13394 14211: contig of 818 bp in length
* 14212 14311: gap of unknown length
* 14312 15465: contig of 1154 bp in length
* 15466 15565: gap of unknown length
* 15566 16380: gap of unknown length
* 16380 17624: contig of 1145 bp in length
* 17625 17724: gap of unknown length
* 17725 18578: contig of 854 bp in length
* 18579 18678: gap of unknown length
* 18679 19774: contig of 1096 bp in length
* 19775 19874: gap of unknown length
* 19875 20700: contig of 826 bp in length
* 20701 20800: gap of unknown length
* 20801 22036: contig of 1236 bp in length
* 22037 22136: gap of unknown length
* 22137 23023: contig of 887 bp in length
* 23024 23123: gap of unknown length
* 23124 24197: contig of 1073 bp in length
* 24197 24297: gap of unknown length
* 24297 25142: contig of 845 bp in length
* 25142 25242: gap of unknown length
* 25242 26385: contig of 1144 bp in length
* 26386 26485: gap of unknown length
* 26486 27325: contig of 840 bp in length
* 27326 27426: gap of unknown length
* 27426 28691: contig of 1266 bp in length
* 28692 28791: gap of unknown length
* 28792 29656: contig of 864 bp in length
* 29656 29755: gap of unknown length
* 29755 30881: contig of 1126 bp in length
* 30882 30981: gap of unknown length
* 30982 31813: contig of 832 bp in length
* 31814 31913: gap of unknown length
* 31914 33057: contig of 1143 bp in length
* 33057 33157: gap of unknown length
* 33157 34008: contig of 851 bp in length
* 34008 34108: gap of unknown length
* 34108 35320: contig of 1213 bp in length
* 35321 35420: gap of unknown length
* 35421 36207: contig of 787 bp in length
* 36208 36307: gap of unknown length
* 36308 37521: contig of 1214 bp in length
* 37522 37621: gap of unknown length
* 37622 38420: contig of 799 bp in length
* 38421 38520: gap of unknown length
* 38521 39726: contig of 1206 bp in length
* 39727 39826: gap of unknown length
* 39827 40635: contig of 809 bp in length
* 40636 40735: gap of unknown length
* 40736 41845: contig of 1110 bp in length
* 41846 41945: gap of unknown length
* 41946 42730: contig of 785 bp in length
* 42731 42830: gap of unknown length
* 42831 43917: contig of 1087 bp in length
* 43918 44017: gap of unknown length
* 44018 44825: contig of 808 bp in length
* 44826 44925: gap of unknown length
* 44926 46114: contig of 1189 bp in length
* 46115 46215: gap of unknown length
* 46215 47031: contig of 817 bp in length
* 47032 47131: gap of unknown length
* 47132 48292: contig of 1161 bp in length
* 48293 48392: gap of unknown length
* 48393 49176: contig of 784 bp in length
* 49177 49276: gap of unknown length
* 49277 50389: contig of 1113 bp in length
* 50390 50489: gap of unknown length
* 50490 51333: contig of 844 bp in length
* 51334 51433: gap of unknown length
* 51434 52765: contig of 1332 bp in length
```

52766	52865: gap of unknown length
52866	53710: contig of 845 bp in length
53711	53810: gap of unknown length
53811	54965: contig of 1155 bp in length
54966	55065: gap of unknown length
55066	55910: contig of 845 bp in length
55911	56010: gap of unknown length
56011	57558: contig of 1548 bp in length
57559	57659: gap of unknown length
57659	58509: contig of 851 bp in length
58510	58609: gap of unknown length
58610	59745: contig of 1136 bp in length
59746	59845: gap of unknown length
59846	60655: contig of 810 bp in length
60656	60755: gap of unknown length
60756	61976: contig of 1221 bp in length
61977	62076: gap of unknown length
62077	62896: contig of 820 bp in length
62897	62997: gap of unknown length
62997	64128: contig of 1133 bp in length
64129	64228: gap of unknown length
64229	65026: contig of 798 bp in length
65027	65126: gap of unknown length
65127	66269: contig of 1143 bp in length
66270	66369: gap of unknown length
66370	67220: contig of 851 bp in length
67221	67320: gap of unknown length
67321	68425: contig of 1105 bp in length
68426	68525: gap of unknown length
68526	69322: contig of 797 bp in length
69323	69422: gap of unknown length
69423	70568: contig of 1146 bp in length
70569	70668: gap of unknown length
70669	71547: contig of 879 bp in length
71548	71647: gap of unknown length
71648	73629: contig of 1982 bp in length
73630	73729: gap of unknown length
73730	74587: contig of 858 bp in length
74588	74687: gap of unknown length
74688	75780: contig of 1093 bp in length
75781	75880: gap of unknown length
75881	77196: contig of 1316 bp in length
77197	77296: gap of unknown length
77297	78651: contig of 1355 bp in length
78652	78751: gap of unknown length
78752	79557: contig of 806 bp in length
79558	79657: gap of unknown length
79658	80774: contig of 1117 bp in length
80775	80874: gap of unknown length
80875	81690: contig of 816 bp in length
81691	81790: gap of unknown length
81791	82928: contig of 1138 bp in length
82929	83028: gap of unknown length
83029	83902: contig of 874 bp in length
83903	84002: gap of unknown length
84003	85193: contig of 1191 bp in length
85194	85293: gap of unknown length
85294	86149: contig of 856 bp in length

Db	93333	CGACTACTAGTAGTACTGCCCCCGGGAAACCTGGACCCCCCGAGCCGAGGCCATTACAC	93392
QY	180	CCGAGGCCCGCGTGTCTGCGCTCCCGCGCGCCAGGTTTAGGCAAGATTCAACCGCTCTTTT	239
Db	93393	CCGAGGCCCGCGCGAGCTTGGCTCCCGCAGCCCGCAGGTTTAGCGGCAAGATCAACCGCTCTTTT	93452
QY	240	TTCTCCGCTACTCTCGGCTACCCCGGGAAACCGCTTGAGCTGGTGGCGCTGATGGCGATT	299
Db	93453	TTCTCCCTACTCTGTGAGTATCCCGAAAACCGCTTGACCTGGTGGCGCTGATGGCGATT	93512
QY	300	CCGTCCTCTCCGACAGGCCCGCGCCCACTCGTGGGCGAGTGTGACGCTCGTACCTTTCG	359
Db	93513	CCGTCCTCTCCGACAGACCCTGTGCCACTCTAGGGCAAAAAGGGAGCGCTCGTACCTTTC	93572
QY	360	CAGGACGCTGTGGAGAGAGGGCGCTGTGTACCGCCCGGTGGAAAGATGGGCTTCC	419
Db	93573	CAGGACACTTCTTTGAGAAAAGGCCCTGTGTGACCCCCCTTGAAAAAATAGCTTTC	93632
QY	420	AGCCGCGGCTTAAAGAGACAGAGAGGGGAGCGTCCGCCCGGACCTGACCAGCGCTGTGGCTT	479
Db	93633	AAACCCCGGGTAAAGAACAAAGAGAGCACTGGCGCC-AGACTGGCCAAACACTGGAGACCT	93691
QY	480	TGCTGAGCTGCGCGCTCATGGG	501
Db	93692	TACTGAACTAGAACTTATAGG	93713

RESULT	10
BV209090	
LOCUS	
DEFINITION	BV209090 456 bp DNA linear STS 02-SEP-2004
ACCESSION	BCI2LI0 3137 Rhesus macaque genomic DNA Macaca mulatta STS genomic clone MfAl137, sequence tagged site.
VERSION	BV209090
KEYWORDS	BV209090..1 GI:51853634
SOURCE	STS.
ORGANISM	Macaca mulatta (rhesus monkey)
AUTHORS	Macaca mulatta Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Macaca. 1 (bases 1 to 456) Spindel,E.R., Pauley,M., Jia,Y., Boyle,N., Jiang,S., Gravett,C., Lupo,S.L., Ali,H., Ojeda,S.R. and Norzgen,R.B. Targeted amplification of the 3' end of rhesus macaque orthologs of human genes Unpublished (2004)
JOURNAL	
COMMENT	

Contact: Spindel ER  
Division of Neuroscience  
Oregon National Primate Research Center  
505 NW 185th Avenue, Beaverton, OR 97006, USA  
Tel: 403-690-5388  
Fax: 503-690-5384  
Email: spindel@ohsu.edu  
Primer A: TTTCACATCTTCGAGAC  
Primer B: CATGTTAAATATCTCAGCA  
STS size: 456

PCR Profile:

Hot Start:	95 degrees C for 2.00 min
Denaturation:	95 degrees C for 0.50 min
Annealing:	51 degrees C for 0.50 min
Polymerization:	72 degrees C for 1.00 min
PCR Cycles:	35
Extension	72 degrees C for 7.0 min
Thermal Cycler:	MJ Instruments PTC100
Protocol:	

Template:	200 ng
Primer:	each 100M
dNTP's:	each 200 uM
Taq Polymerase:	0.05 units/ul (Fast Start High Fidelity, Roche)
Total Vol:	50 ul

## Buffer:

MgCl<sub>2</sub>: 1.8 mM  
Fast Start polymerase reaction buffer (Roche)

Bases 1-321 are 95% homologous (Blast) to bases 566-886 of NM\_020396.2. Primers were chosen to amplify genomic DNA in the 3' region of BCL2L10. As human sequence was used to design the primers, the primer sequences are not included in the rhesus sequence provided below. To obtain additional information regarding primers or clones contact: Dr. Robert Norgren; Dept of Genetics Cell Biology & Anatomy; University of Nebraska Medical Center; 966395 Nebraska Medical Center; Omaha, NE 68198. Email: rnorgren@unmc.edu

A database containing sequences associated with this project can be found at: <http://rhesusgenecchip.unomaha.edu/index.html>.

## FEATURES

## SOURCE

1..456  
/organism="Macaca mulatta"  
/mol\_type="genomic DNA"  
/strain="Indian origin"  
/db\_xref="taxon:9544"  
/clone="MMA3137"  
/clone\_lib="Rhesus macaque genomic DNA"  
/dev\_stage="Adult"  
/note="Organ: Liver; Vector: pGEM-T Easy; V-type: Plasmid; cDNA amplified from rhesus genomic DNA with the human forward and reverse primers listed above and subcloned into pGEM-T Easy."  
1..456  
/gene="BCL2L10"  
/note="BCL2-like 10"  
<1..>456  
/gene="BCL2L10"

## STS

## ORIGIN

Query Match 33.3%; Score 295.4; DB 11; Length 456;  
Best Local Similarity 95.0%; Pred. No. 2.2e-41;  
Matches 305; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 566 CCGTTTCCAGTGGCTTTTGGAGAAACAGTGGTCCAGGCTTTCTGTCAAGCTTTGTA 625  
Db 1 CCGTTTCCAGTGGCTTTTGGAGAAACAGTGGTCCAGGCTTTCTGTCAAGCTTTGTA 60

QY 626 ACAACAGGCTTCATTATCTCTGACAGATTATTAGTTTAACTTTTAAACCGC 685  
Db 61 GCAACAGGCTTCGTTATCTCTGACAGATTATTAGTTTAACTTTTAAACCGC 120

QY 686 TTCTACCTGCGCACTGTGACCACTAATGACAGATGTGAGAAACAAGACTGAGGA 745  
Db 121 TTCTACCTGCGCACTGTGACCACTAATGACAGATGTGAGAAACAAGACTGAGGA 180

QY 746 AAGCAGCTTCCCGCCAGCCAGAGCTTTTATCTGAATGATCAAGAGTCCCTGAGGTGG 805  
Db 181 AAGCAGCTTCCCGCCAGCCAGAGCTTTTATCTGAATGATCAAGAGTCCCTGAGGTGG 240

QY 806 TGATTTGGCCAGTGTGTTTAACTTGTGACAAGTACTCAGGTGTGAGAGACAAGATCAAT 865  
Db 241 TGATTTGGCCAGTGTGTTTAACTTGTGACAAGTACTCAGGTGTGAGAGACAAGATCAAT 300

QY 866 GGCTCTTCTGAGTGAAGA 886  
Db 301 GGCTCTTCTGAGTGAAGA 321

## RESULT 11

AF102501 AF102501 1209 bp mRNA linear ROD 21-JAN-1999  
LOCUS Mus musculus Bcl-2 homolog (Bcl-2) mRNA, complete cds.  
DEFINITION  
ACCESSION AF102501  
VERSION AF102501.1 GI:4165137  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
1 (bases 1 to 1209)  
AUTHORS Song,Q., Kuang,Y., Dixit,V.M. and Vincenz,C.  
TITLE Bcl-2, a novel negative regulator of cell death, interacts with

JoRNAL EMBO J. 18 (1), 167-178 (1999)  
MEDLINE 99094902  
PUBMED 9878060  
REFERENCE 2 (bases 1 to 1209)  
AUTHORS Song,Q.Z., Kuang,Y.P., Dixit,V.M. and Vincenz,C.  
TITLE Direct Submission  
JOURNAL Submitted (28-Oct-1998) Pathology, University of Michigan, 1301 Catherine Road, Ann Arbor, MI 48109, USA

## FEATURES

## SOURCE

1..1209  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6D2 F1/J"  
/db\_xref="taxon:10090"  
/chromosome="9"  
/map="D9"  
/tissue\_type="ovary"  
1..1209  
/gene="Bcl-2"  
117..692  
/gene="Bcl-2"  
/product="Bcl-2 homolog"  
/protein\_id="AA08703.1"  
/db\_xref="gi:4165138"  
/translation="MADSDPILHRTRLSDYIFPCAREDPTEPPSYEALNLS  
VTROIQEHQEFSSFCESRGNRLIELYKQAMADKLSKDQFSMLVLAFASTLMN  
QGPYAVKOKRDLGNRVIVTRDCLIVNPLVNLGRRHRRLBALGMDQFCRPFKN  
PLPLGPRRLILQFLSGFRATLPIFMKL"

## gene

## CDS

## ORIGIN

Query Match 29.2%; Score 259.2; DB 10; Length 1209;  
Best Local Similarity 63.1%; Pred. No. 3.8e-35;  
Matches 477; Conservative 0; Mismatches 253; Indels 26; Gaps 4;

QY 86 GACCCGCTGCGGAGACCGCAGAGCTGTGCTGCGCCGACTACCTGCGGCTCGCCG 145  
Db 132 GACCCGCTGCGGAGACCGCAGAGCTGTGCTGCGCCGACTACCTGCGGCTCGCCG 191

QY 146 GAACCCGCGACCCCGAGCGCGGCGCATCCAGCCCGAGCGCGCGTGGCTCGCGCG 205  
Db 192 GAACCCGCGACCCCGAGCGCGGCGCATCCAGCCCGAGCGCGCGTGGCTCGCGCG 251

QY 206 GCCGCAAGTTACGCGCAGATTCAACCGCTCTTTTCTCCGCTTACCTGCGCTACCCGCG 265  
Db 252 ACTAGGAGAGATCCAGAGAGACCAAGATTTTCTCTCTCTGCGGAAAGCCGCGCG 311

QY 266 AACCGCTTGAAGCTGTGCGCGCTGATGCGGATTCGCTCTCTCGACAGCCCGCGCG 325  
Db 312 AATGCGCTGAGCTGTGCGGAAACAGATGCAATAGTGTCTCTCCAAAGACCAAGCTTC 371

QY 326 AACTGCGGCAAGTGTGAGCGCTGTGACCTTCCGAGGAGCGTGTGAGAGAGAGCGCG 385  
Db 372 AACTGCGGCAAGTGTGATCTCTGCGCTTCCGAGGAGCGCTTATATACAGAGCGCT 431

QY 386 CTGTGACCGCCCGGTGAGAGAGTGGGCTTCCAGCCCGCTTAAAGAGACAGAGGCG 445  
Db 432 TACATGGCTGTCAAGCAGAGAA-----GGATCTGGGGAATGTGTG 473

QY 446 GACGTGCGCGGAGCTGCGAGCGCTGTGCTGCTGAGCTGCGGCTCATGGG--- 502  
Db 474 ATATGACCCCGAGCTGTCTCTCATATGTGAATCTTGTAAATCTGCTATGGGCGGT 533

QY 503 CAGACCGCGCGCTGCGTCAAGCTCAAGGCGCGCTGAGATGAGCTTTTGTCACTTCTTCAAG 562  
Db 534 CAGACCGCGCGCTGAGAGCTCTCGCGCGCTGAGAGAGCTTTTGTCCGCTTTCAAG 593

Qy 563 ACCCCCTTTCAGCTGGCTTTTGGAGAAAACAGCTGCTCCAGCTTTTCTGTCAAGCTTG 622  
 Db 594 AATCTTTACCGCTCGGCTTCTGGAGAAAGATTGATTCAGGCTTTTCTGTCAAGCTTC 653  
 Qy 623 TTACCAACAGCTTCATTATCTCTGACAGCATTAATATGAGTTTAAACTTTTAAACC 682  
 Db 654 TTTCGCAAGCAGCATCTTTTATCTGGAACCTTTAT---AAGTTTAAATTTTAAAG 709  
 Qy 683 CGCTTCACCTGCCCCAGCTGACCACTAAATGACAGTGTGTGAGAACAAAGACTAG 742  
 Db 710 CGATCTTACCTACCTACCTGAGACCTCTTAAAGAACTTGGGTATTTGATTAAGAA 769  
 Qy 743 GAAAGACCTTCCCAAGCCCAAGCTTTTATCTGATGATGATCAAGAGTCTGAGG 802  
 Db 770 CTGAGGAAGTCTCCAGGCTTAGAGACATTTCTACCTGATGC-TACATGAGATTCTGGGT 828  
 Qy 803 TGTGATTTGGCCAGCTGTTTATCTTGTGACAAAGTA 838  
 Db 829 TTGTATGGGGCCAGCATTTGAAAGGGGCCCATGTA 864

## RESULT 12

AF067660 1225 bp mRNA linear ROD 04-DEC-1998  
 LOCUS Mus musculus Bcl-2 homolog (Diva) complete cds.  
 AF067660  
 VERSION AF067660.1 GI:3955265  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 1225)  
 Garcia, I., Koseki, T., Hu, Y., Chen, S. and Nunez, G.  
 Diva, a Bcl-2 homologue that binds directly to Apaf-1 and induces  
 BH3-independent cell death  
 J. Biol. Chem. 273 (49), 32479-32486 (1998)

## JOURNAL

MEDLINE 99047617  
 PUBMED 9829980  
 2 (bases 1 to 1225)  
 Inohara, N. and Nunez, G.  
 Direct Submission  
 Submitted (20-MAY-1998) Department of Pathology, University of  
 Michigan Medical School, 1500 E. Medical Dr., Ann Arbor, MI 48109,  
 USA

## FEATURES

## Source

Location/Qualifiers  
 1..1225  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="B6D2F1/J"  
 /db\_xref="taxon:10090"  
 1..1225  
 /gene="Diva"  
 143..718  
 /gene="Diva"  
 /note="Diva"  
 /codon\_start=1  
 /product="Bcl-2 homolog"  
 /protein\_id="AAC83150.1"  
 /db\_xref="gi:3955266"  
 /translation="MADSODPLHRTRLISDYIFPCAREPTPEPTTSEALILS  
 VTRIOEHOFSSFCESRKNRLLELYQAKDLISKQDPSQQLVLLAPATLW  
 QGFYMAVKQKDLGNRVIVTRDCCILVFLNLMGRHRARLALGWDGFCRFFKN  
 PLPLGFWRRLRIQAFLSGFATALFIWKRL"  
 1197..1202  
 /gene="Diva"  
 1225  
 /gene="Diva"

## CDS

polyA\_signal  
 polyA\_site  
 1225  
 /gene="Diva"

## ORIGIN

Query Match 29.2%; Score 259.2; DB 10; Length 1225;  
 Best Local Similarity 63.1%; Pred. No. 3.8e-35;

Matches 477; Conservative 0; Mismatches 253; Indels 26; Gaps 4;  
 Qy 86 GACCCGTCGGAGAGCCAGAGCTGTGGCCGCTACCTGGGGTACTGGCCCGG 145  
 Db 158 GACCCAGTCAGTAAAGCAGTAAAGCGCTGCTGTCTCAATATTTCTTGGCAGG 217  
 Qy 146 GAACCCGAGACCCCGAGCCGCGCAATCCAGCCGAGCCGCGCTGCTGCGG 205  
 Db 218 GAGCCGAGACCCAGAGCCAGCCGCGCAATCTGTGAGAGGGGCTTCTGCTGTG 277  
 Qy 206 GCCCGAGTTACCGAGATTACCGGCTCTTTTCTCCGCTACCTCGGCTACCCG 265  
 Db 278 ACTAGGAGATTCAGAGAGACCAAGAAATTTTCTCTCTCTGTGGAAGCCGAGG 337  
 Qy 266 AACCGCTTCAGCTGTGGCCCTGATGGCGGATTCGCTCTCTCCAGAGCCGCGCC 325  
 Db 338 AATGCGCTGAGCTGTGTAAACAGATGCAATATGTTGCTCTCCAAAGACCAACTTC 397  
 Qy 326 ACTGGGCGAGAGTGTAGCTGCTGCTTCCAGGAGAGCTGCTGAGAGAGGCGG 385  
 Db 398 AGCTGAGCCAACTGTGTATCTCTGAGCTTCCGCGGAGCGCTTATGAAATCAAGCCCT 457  
 Qy 386 CTGTGACCGCCGCTGTGAGAGAGTGGGCTTCCAGCCGCGCTAAAGAGCAGAGGCG 445  
 Db 458 TACATGGCTGTCAAGCAGAGAA-----GGATTTGGGGAATCTGTCTC 499  
 Qy 446 GACGTGCGCCGGAGCTGCGAGCGCTGTGTGCTTGTGAGCTGCGGCTATGGGG-- 502  
 Db 500 ATATGAGACCCAGACTGTCTCTCATGTGAACTTCTGTATTAATCTGCTATAGGGGGT 559  
 Qy 503 CAGACCGCGCTGAGCTGAGAGCTCAGGCGCGCTGGAGTGGCTTTTGTCACTTTCAGG 562  
 Db 560 CGGACCGCGCGAGGCTGTGAGAGCTCTGCGCGCGCTGGAGTGGCTTTTGTCACTTTCAGG 619  
 Qy 563 ACCCCCTTTCAGCTGGCTTTTGGAGAAACAGCTGTGCTGAGCTTTTCTGTCACTTTC 622  
 Db 620 AATCTTTACCGCTGTGGCTTGTGAGAAATGTGATTCAGGCTTTTCTGTCACTTTC 679  
 Qy 623 TTACCAACAGCTTCATTATCTCTGAGACAGATTAATATGATTTTAAACTTTTAAACC 682  
 Db 680 TTGCAACAGCAGCATCTTTTATCTGAAAGAGTTAT---AAGTTTAAATTTTAAAG 735  
 Qy 683 CGCTTTCACCTGAGCACTGAGCACTAAATGACAGATGTGAGAAAGAAAGTGG 742  
 Db 736 CGATCTTACCTACCTACCTGAGAACCTCTTAAAGAAATTTCTGGTAATTTGATTAAGAA 795  
 Qy 743 GAAAGACACTTCCCAAGCCCAAGCTTTTATCTGATGATCATACAGAGTCTGAGG 802  
 Db 796 CTGAGGAAGTCTCCAGGCTTAGAGACATTTCTACCTGATGC-TACATGAGATTCTGGGT 854  
 Qy 803 TGTGATTTGGCCAGCTGTTTATCTTGTGACAAAGTA 838  
 Db 855 TTGTATGGGGCCAGCATTTGAAAGGGGCCCATGTA 890

## RESULT 13

BC052690 1257 bp mRNA linear ROD 08-OCT-2003  
 LOCUS Mus musculus Bcl2-like 10, mRNA (cDNA clone MGC:60542  
 IMAGE:30052580), complete cds.  
 ACCESSION BC052690.1 GI:30851238  
 VERSION  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 1257)  
 Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,  
 Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

JOURNAL	TITLE	REMARK
JOURNAL	STRAPLETON, M., Soares, M.B., Donald, M.F., Casavant, T.L., Schnee, T.E., Brownstein, M.J., Ueding, T.B., Toohilly, S., Carninci, F., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., Mobaw, P.J., McMernan, K.U., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Holly, S.W., Vallalon, D.K., Muzny, D., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalka, U., Smalius, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.	
JOURNAL	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
PUBMED	2238257	
REFERENCE	12477932	
AUTHORS	2 (bases 1 to 1257) Strausberg, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-MAY-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	
COMMENT	Contact: MGC help desk	

cDNA Library Preparation: Yulan Piao and Minoru Ko (National Institute on Aging, NIH: <http://193sn.grc.nia.nih.gov/cDNA/>)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (NLN)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
Akhter, N., Ayale, K., Beckertow-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hudson, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Madsen, O.L., Mastello, C., Mackeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Startitop, S., Thomas, P.J., Touchman, J.W., Tsougeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLAN at: <http://image.llnl.gov>  
 Series: IRAX Plate: 112 Row: h Column: 18  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7304926.  
 Location/Qualifiers

```

source
1..1257
/organism="Mus musculus"
/mol_type="RNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="WGC:60542 IMAGE:30052580"
/tissue_type="Egg, unfertilized, mouse"
(clone_lib="NIA Mouse Unfertilized Egg cDNA Library
Long")
/lab_host="DH10B"
/notes="vector: pSPORT1"
1..1257
/gene="Bcl2l10"
/notes="synonyms: Boo, Diva"
/db_xref="LOCusID:12049"
/db_xref="MGI:1330841"
145..720
/gene="Bcl2l10"
/codon_start=1
/product="Bcl2l10 protein"
/protein_id="AAH52690.1"

```

```
misc_feature
/db_xref="GI:30851239"
/db_xref="LocusID:12049"
/db_xref="WGI:1330841"
/translation="MADSDQPLHGRTRRLSDYIFPCAREPDTPPEPTSEVALNNS
VTRR10QH0HFSSFCESGRNLELVKQAMDKLSKQDFSMQLVMLAFATLANN
QGPYMAVKQKRDLDGNRYIVTRDCLIVNPLYNLLMGRHRALEALGSMGFCFFKN
PLPLGFWRLDLOFLSGFPATLFIWKRL"
268..597
/gene="Bcl2l10"
/note="BCL: Region: BCL (B-Cell Lymphoma)"
/db_xref="CCD:smart00337"
```

ORIGIN			
Query Match	29.0%	Score 257.6	DB 10; Length 1257;
Best Local Similarity	63.0%	Pred. No. 7.1e-35;	
Matches 476; Conservative	0;	Mismatches 254;	Indels 26; Gaps 4;

QY	86	GACCCTGTCGGGAGCGCAGCACTGTTGTGTGGCCGACATCACTCGGGGTATCTGGCCCGG	145
Db	160	GACCCTACTGATGGACCGCATAGACGGGTGTGTCTGATTAATATTTCTTCTGGGACCG	219
QY	146	GAATCCCGGCACATCCCGGAGCCGGGCCCATTCACGCCCCGAGGCGCGCGTGTCCGCTCCGGC	205
Db	220	GAGCCGGAACAACCCAGAGGCCACCGCCACGTCTGTCAAGCGGCGCTTGTCCGTGTGTG	279
QY	206	GCCGCAAGTTACGGCAGATTCACCGGTCCTTTTCTCCGCTACTCTCGGCTACCCCGGG	265
Db	280	ACTGAGCAGATTCACGACGAGACACCAAGAAATTTTTTCTCTTCTCTGCGAAAGCCGGGGC	339
QY	266	AAACCGCTTGAAGTGTGGCGCTGATGGCGGATTTCCGTGTCTTCGACAGCCCCGAGCCCC	325
Db	340	AATGCGCTCGAGCTGTGTGAACACATATGGCAGATTAATGTGCTCTCCAAAGACCAAGACTTC	399
QY	326	ACCTGGGGCAGAGGTGTGACGCTGTGATCACTTGGCAGAGGACGCTGCTGAGAGAGGGCCG	385
Db	400	AGCTGAGGCCAATGTGTGATGTCTCTGACCTTCGCGGGAGGCTTAATGAATCAAGGCCCT	459
QY	386	CTGTGTACCCGCCGTGTGAAGAAAGTGGGAGCTTCCAGCGCGGCTTAAAGAGACAGAGAGGC	445
Db	460	TACATGGCTGTCAAGCAGAGA-----GGATTTGGGGAAATCGTGTCTC	501
QY	446	GACGTGCGCCGGGACTGCAAGCGCTGTGTGGCTTGTCTGAGCTGCGGCTCATGTGGG---	502
Db	502	ATATGTGACCCCAAGACTGTGTCTATATGTGAACCTTTCTATATATCTGCTCATGTGGGCGT	561
QY	503	CAGACCGCGGCTGTGGCTGTGACAGGCGCGGTGGAGTGGCTTTTGTATCACTTTTCAGG	562
Db	562	CGGACCGGCGCAGGCTGTGAGGCTCTCGGCGGCTGGGATGTCTTTTGTCCGCTTCTTCAAG	621
QY	563	ACCCCTTTCACGTGGCTTTTGTGAGAAAACAGCTGTGTCAAGCTTTTCTGTCAATGCTTG	622
Db	622	AATCTTTTACCGCTCGGCTTCTGTGAGAAAGATTCGTGATTCAGGCTTTTCTGTCAAGCTTCTC	681
QY	623	TTTAACAACAGCTTCATTATCTCTGTGACACGATTAATTAGATGTTTAAAACTTTTAAC	682
Db	682	TTTGTGAAGGCATCTTTTATCTGTGAAGAGTTAT---AAGTTTAAAAATTTTAAAG	737
QY	683	CGCTTCTACCTGCCCACTGTGTGCCAACATAATGACAGATGTGTGAGAACAAAGACTGAG	742
Db	738	CGATCTTACTACTTACTGTGAACCCCTTAAGGAACAATTCGTGGTAAATTGGATTAAGAA	797
QY	743	GGAAAGCATTTCCTCCCAACCCAGACGTTTATCTGTGAATGCATACAGAGATCTGTAGG	802
Db	798	CTGAGGAAGTCTCTCAGGCTAGAGACATTTCTACTCTGATGTC-TACATGTGGATTCGTGGT	856
QY	803	TGTGATTTTGGCGCAGTGTTTTAATCTGTGACAACTA	838
Db	857	TTGTGATGTGGCGCAGCAATTGAAAAGGGGCCCATGTA	892

RESULT 14		
AY029163		
LOCUS		
AY029163	1074 bp	mRNA linear ROD 19-JUN-2003



DEFINITION Rattus norvegicus BCL2L10 (Bcl2l10) mRNA, complete cds.  
 ACCESSION AY029163  
 VERSION AY029163.1 GI:13641257  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE Rattus norvegicus  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 REFERENCE 1 (bases 1 to 1074)  
 AUTHORS Itoh, T., Itoh, A. and Pleasure, D.  
 TITLE Bcl-2-related protein family gene expression during oligodendroglial differentiation  
 JOURNAL J. Neurochem. 85 (6), 1500-1512 (2003)  
 MEDLINE 22672518  
 PUBMED 12787069  
 REFERENCE 2 (bases 1 to 1074)  
 AUTHORS Itoh, T., Itoh, A. and Pleasure, D.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-MAR-2001) Neurology Research, The Children's Hospital of Philadelphia, Abramson Research Center, Room 516 I, 3517 Civic Center Boulevard, Philadelphia, PA 19104, USA  
 FEATURES  
 source  
 1..1074  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 1..1074  
 /gene="Bcl2l10"  
 12..569  
 /gene="Bcl2l10"  
 /note="Bcl2l10"  
 /product="Bcl2L10"  
 /codon\_start=1  
 /protein\_id="AAK31792.1"  
 /db\_xref="GI:13641258"  
 /translation="MGPDIODRTRRLTDVILPCARAPNTEPELPTSEVALLRSSVTS QIQEHODLFRSPDYGNRLIELYTMADSLSDSEFNRMLYMLAIPVGLTMRQR TVRRKDRRLRLRDCVYLIVSLVYRNLGTRHRSMLRAGMDGFCQFFKNPLPGR WRLRLRALISCFPATAIYIMKCL"  
 ORIGIN  
 Query Match 27.5%; Score 244.2; DB 10; Length 1074;  
 Best Local Similarity 61.7%; Pred. No. 1.5e-32;  
 Matches 455; Conservative 0; Mismatches 253; Indels 29; Gaps 3;  
 QY 80 ATGAGCCGACCGCTGCGGAGACGACGAGCTGTGCGGCGACTAGCTGGGTACTGC 139  
 DB 12 ATGGGTGACCCGCTGCGAGATCGCATAGACGCTGCTGACTGACTATATTGTTCTGC 71  
 QY 140 GCCCGGAAACCCGACCCCGGACCGCGCCGATCCAGCCCGAGCCGCGTCTGCGC 199  
 DB 72 GCAAGGGGCGCGAACCCTCGACCTGCGCCAGCTGTGTTGAGGCGGCGCTTCTGCGC 131  
 QY 200 TCCGCGCGCGCGAGTTCACGACATTCACCGGTCTTTTCTCCGCTACTCTGGGCTAC 259  
 DB 132 TCTGTGCTAGTGCATTCACAGAGACACAGATCTTTTCAACTCTTCCGCGACTAC 191  
 QY 260 CCCGGGAAACGCTTGCAGCTGTGCGCTGATGCGGATTCGCTCTCCGACAGCCCC 319  
 DB 192 CAGGGCAACCGCTGAGAGCTGTGACACAGATGCGGATGAGTGTCTTCCAAAGACCA 251  
 QY 320 GGGCCCACTGAGGAGAGTGTGACGCTGTGACCTTTCGAGGAGCGCTCTGAGAGA 379  
 DB 252 GAGTTCAACTGGGCGCGCTGTGATGCTCTGCGCTTCTGAGGAGCGCTATGAAACCA 311  
 QY 380 GGGCGCGTGTGACCGCCCGGTGAAGAAGTGGGCTTCCAGCCCGGCTAAAGAGCAG 439  
 DB 312 GACAGAGCTGTTAAGCGAGAGGAGATCAAGAAACCGTCTTCTAC----- 357  
 QY 440 GAGGGACGCTGCGCGGAGCTGCAAGGCGCTGTGCTTGTGAGCTCGGAGCTCATG 499

DB 358 -----TGAGCAGACTGCTATCTCATGTGAGCTTGCTGTAACAATGCACTGACA 407  
 QY 500 GGGAGACACCGCGCTGCTGCTGACAGCTTCAGGCGCGCTGGATGCTTTTGCACTTCTC 559  
 DB 408 GAGCGGATGCTGCTGCTGAGAGCTTCACGCTGCGGAGCTTTTGCAATTTCTC 467  
 QY 560 AGAACCCCTTTCACAGCTGCTTTTGAAGAAACAGCTGCTGCTGCTTTTCTGTCATGC 619  
 DB 468 AAGAACCCCTTACACACCGGCTTCTGAGAGAGATTGCTGATCCGGCTATTTCTCTCT 527  
 QY 620 TTGTTAACACAGCTTCAATTATCTCTGACACAGATTATTAAGTTTAAACTTTTA 679  
 DB 528 TTTCTTCAACGCGCATCTTTATATCTGAGAAAGTTTAT---AAGTTTAAATTTTAA 583  
 QY 680 ACCGCTTACACCTGCGCCCACTGACCACTAATAGATGTGAGAACAACT 739  
 DB 584 AATGATCTTACCTGCTTAACTGCGGCGCTTAAAGGACAACTCTGGTATGTTGATAA 643  
 QY 740 GAGGAAAGCAGCTTCCCGCACCCAGACGTTTATCTGATGATGATCAAGAGTCTG 799  
 DB 644 GAATGAGAAATCTCTGCTGCTAGACATTTCTACTGATGC-TACACGAGTTCCG 702  
 QY 800 AGGTGTGATTGGCCA 816  
 DB 703 GATTGGAGTGGCCTA 719  
 RESULT 15  
 AC133947 168997 bp DNA linear ROD 08-SEP-2004  
 LOCUS Mus musculus chromosome 9 clone RP24-18912, complete sequence.  
 DEFINITION AC133947  
 ACCESSION AC133947  
 VERSION AC133947.3 GI:51921460  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 168997)  
 AUTHORS Wilson, R.K.  
 TITLE The sequence of Mus musculus clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 168997)  
 AUTHORS McPherson, J.D. and Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 3 (bases 1 to 168997)  
 AUTHORS Wilson, R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-JUL-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 4 (bases 1 to 168997)  
 AUTHORS Wilson, R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-SEP-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 COMMENT On Sep 8, 2004 this sequence version replaced gi:50199129.  
 ----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu  
 Contact: submissions@watson.wustl.edu  
 Project Information  
 Center project name: M.BB0189102  
 -----  
 FEATURES  
 source  
 1..168997  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="9"

ORIGIN

/clone="RP24-18912"

Query Match 16.6%; Score 147; DB 10; Length 168997;

Best Local Similarity 61.5%; Pred. No. 7.2e-16;

Matches 281; Conservative 0; Mismatches 155; Indels 21; Gaps 2;

```
QY      86 GACCCGCTGCGGAGACCGACCGAGCTGTTGCTGCGCGACTACCTGAGGTAATGCGCCCG 145
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166427 GACCCACTGCAATGACGACGAGGCTGCTGCTGTAATATCTTCTGCGCACG 166486
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      146 GAACCCGCGACCCCGGACCGGCGCATCCAGCCCGAGGCCCGCTGCGCTCCGCG 205
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166487 GAGCCGAGACCCCGACCGACCCCGCCACGCTGTCAGGCGGCGCTTGTGCTGTG 166546
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      206 GCCGCGAGTTACGCGAGATTACCGGTCCTTTTCTCGGCTTACCTGCGCTACCCCGG 265
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166547 ACTAGGCGATTCAGCAGAGAGACCAAGATTTTCTCTCTGCGAAGCCCGGCGC 166606
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      266 AACGCTTCAGAGCTGAGCGCTGATGCGGATTCCTGCTCTCCGACAGCCCGGCCC 325
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166607 AATGCTCTGAGCTGTGTAACAGATGCGAGTAAGTTGCTCTCCAAAGACCAAGACTTC 166666
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      326 ACTGCGGCGAGAGTGTGACGCTGTAACCTTCGAGGAGCGCTGTGAGAGAGGCGCG 385
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166667 AGCTGAGCCCACTGTGTATGCTCTGCGCTTCGCGGAGCGCTTATGATCAAGGCCCT 166726
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      386 CTGCTGACCGCCCGGTGAGAGAGTGGGCTTCCAGCCGCGCTAAAGAGAGAGAGGAGC 445
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166727 TACATGCTGTCAAGCAGAAAG-----GGATCTGGGATATGCTGTC 166768
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      446 GACGTGCGCCGCGAGCTGCGCAGCGCTGTGAGCTTGTGAGCTCGCGCTCATGCGG--- 502
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166769 ATAAGTACCCGAGACTGCTGTCTCATAGTGAACCTTCTGTATTAATCTGCTCATGGGCGT 166828
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      503 CAGCACCGCGCTGCTGCTGAGGCTCAGGCGCGCTGCG 539
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166829 CGGCACCGCGCGAGGCTCTCGCGCGCTGCG 166865
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: June 7, 2005, 21:57:49  
Job time : 6833.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 08:53:44 ; Search time 867,812 Seconds  
(without alignment)  
6050,630 Million cell updates/sec

Title: US-10-071-174A-1

Perfect score: 887

Sequence: 1 cggggccaaagaacaccgca.....ctctccttgagcgaagaa 887

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: N\_Geneseq\_16Dec04:\*  
2: geneseqn1980s:\*  
3: geneseqn1990s:\*  
4: geneseqn2000s:\*  
5: geneseqn2001as:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	887	100.0	887	6	AAD46683 Human Bcl
2	812.4	91.6	1168	3	AAZ90049 Hydrophob
3	582	65.6	582	3	AAZ90039 Hydrophob
4	506.8	57.1	522	4	AAH47022 Human Bcl
5	457.2	51.5	548	4	ABK41913 CDNA enco
6	457.2	51.5	548	9	ADB59580 Connectiv
7	324.6	36.6	874	6	ABQ44404 Oligonuc
8	324.6	36.6	874	6	ABQ44405 Oligonuc
9	320	36.1	874	6	ABQ44403 Oligonuc
10	320	36.1	874	6	ABQ44402 Oligonuc
11	240.2	27.1	650	10	ADD34132 Mouse mit
12	154	17.4	835	10	ADD34582 Mouse mit
13	88.6	10.0	338	10	ADD34583 Mouse mit
14	68	7.7	2000	8	ADA71938 Rice gene
15	60	6.8	60	6	ABN42801 Human spl
16	59.4	6.7	114955	2	AAK53491 Human ade
17	57.4	6.5	9222	8	AAI61171 Actinosyn
18	57.4	6.5	39949	12	ADF31997 Full leng
19	57.4	6.5	48200	12	ADF31998 Coemid 2A
20	57.4	6.5	82746	8	AAI61224 Actinosyn

21	56	6.3	536	10	ADB68842	ADB68842 Minority
22	55.6	6.3	1765	5	AA883822	AA883822 DNA encod
23	55.6	6.3	1956	5	AA573207	AA573207 DNA encod
24	55.6	6.3	1956	5	AA583791	AA583791 DNA encod
25	55.6	6.3	3808	10	AD121903	AD121903 Novel hu
26	55.6	6.3	4023	13	ADRO6595	ADRO6595 Full len
27	55.2	6.2	3158	6	AA150615	AA150615 Human glt
28	52.4	5.9	9975	8	AAI61173	AAI61173 Actinosyn
29	51.8	5.8	113193	8	ADG98445	ADG98445 Streptom
30	51.4	5.8	1293	12	ADG56245	ADG58245 Acyl-spe
31	51.4	5.8	73882	13	ADG73531	ADG7351 tcp gene
32	51	5.7	58857	3	AAAS8471	AAAS8471 Nucleotid
33	50.6	5.7	1437	8	ADA69472	ADA69472 Rice gene
34	50.6	5.7	14055	8	AAI61170	AAI61170 Actinosyn
35	50.4	5.7	114955	2	AAK53491	AAK53491 Human ade
36	50.2	5.7	834	10	ADK57446	ADK57446 Plant DNA
37	49.8	5.6	1194	8	ABX56038	ABX56038 M. echino
38	49.4	5.6	3540	12	ADOB85407	ADOB85407 Streptom
39	49.4	5.6	4440	12	ADOB85409	ADOB85409 Streptom
40	49.4	5.6	4460	12	ADOB85408	ADOB85408 Streptom
41	49.2	5.5	6200	8	ACA27156	ACA27156 Prokaryot
42	49	5.5	11100	10	AD123891	AD123891 Streptom
43	49	5.5	13315	10	AD123887	AD123887 Streptom
44	48.8	5.5	27541	4	AD17185	AD17185 Streptom
45	48.8	5.5	125401	4	AD17186	AD17186 Streptom

ALIGNMENTS

# ALIGNMENTS

RESULT 1	AAZ90049	Hydrophob
ID	AAD46683	standard; DNA; 887 BP.
XX	AAZ90049	Hydrophob
XX	AAZ90039	Hydrophob
XX	AAH47022	Human Bcl
XX	ABK41913	CDNA enco
XX	ADB59580	Connectiv
XX	ABQ44404	Oligonuc
XX	ABQ44405	Oligonuc
XX	ABQ44403	Oligonuc
XX	ABQ44402	Oligonuc
XX	ADD34132	Mouse mit
XX	ADD34582	Mouse mit
XX	ADD34583	Mouse mit
XX	ADA71938	Rice gene
XX	ABN42801	Human spl
XX	AAK53491	Human ade
XX	AAI61171	Actinosyn
XX	ADF31997	Full leng
XX	ADF31998	Coemid 2A
XX	AAI61224	Actinosyn
PT	New isolated or recombinant Bcl-B nucleic acids and polypeptides, for	



CC activity; haemostatic and thrombolytic activity (e.g. treating  
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and  
CC tumour inhibition activity. The polynucleotides are also stated to be  
CC useful for gene therapy. Other activities include inhibiting infections  
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,  
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,  
CC skin, effecting biohythms or cardiac cycles; enhancing fertility;  
CC treatment of depression; treatment of pain; hormonal or endocrine  
CC activity. The polynucleotides may also be used for recombinant expression  
CC of the protein

XX Sequence 1168 BP; 276 A; 312 C; 300 G; 280 T; 0 U; 0 Other;

Query Match 91.6%; Score 812.4; DB 3; Length 1168;

Best Local Similarity 99.9%; Pred. No. 7.1e-173;

Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

74 ACCACCATGCGCGACCGCTCGGGAGCGCACCGAGCTGTGTGCGCGACCTACCTGAGG 133

1 ACCACCATGCGCGACCGCTCGGGAGCGCACCGAGCTGTGTGCGCGACCTACCTGAGG 60

134 TACTGCGCGCGGAAACCGGACCGCGCGCGCATCCAGCCCGAGCGCGCGT 193

61 TACTGCGCGCGGAAACCGGACCGCGCGCGCATCCAGCCCGAGCGCGCGT 120

194 CTGCGCTCCCGCGCGCGCGGATACCGGAGATCAACCGGCTCTTTCTCCGCGTACCTC 253

121 CTGCGCTCCCGCGCGCGCGGATACCGGAGATCAACCGGCTCTTTCTCCGCGTACCTC 180

254 GGTACACCGCGGAAACCGCTTCGAGCTGTGTGCGCGATGCGGATTCCTCGCTCCGAC 313

181 GGTACACCGCGGAAACCGCTTCGAGCTGTGTGCGCGATGCGGATTCCTCGCTCCGAC 240

314 AGCCCGCGCGCGACCTCGGGAGCGAGTGTGACGCTCGTACCTTCGACGAGACGCTGCTG 373

241 AGCCCGCGCGCGACCTCGGGAGCGAGTGTGACGCTCGTACCTTCGACGAGACGCTGCTG 300

374 GAGAGAGGCGCGCTGTGACCGCGCGCGGAAAGTGGGCGCTTCACGCGCGGCTAAG 433

301 GAGAGAGGCGCGCTGTGACCGCGCGCGGAAAGTGGGCGCTTCACGCGCGGCTAAG 360

434 GAGAGAGGCGCGAGTGTGCGCGCGGAGCTGCGAGCGCTGTGTGCTGAGCTCGCGG 493

361 GAGAGAGGCGCGAGTGTGCGCGCGGAGCTGCGAGCGCTGTGTGCTGAGCTCGCGG 420

494 CTATGCGGCGAGCACCGCGCTGTGACGCTGACGCGCGCTGGGATGCTTTGTAC 553

421 CTATGCGGCGAGCACCGCGCTGTGACGCTGACGCGCGCTGGGATGCTTTGTAC 480

554 TTCTTCAGGACCGCGCTTCACCTGCGCTTTTGGAGAAACAGCTGTGCGAGCTTTCTG 613

481 TTCTTCAGGACCGCGCTTCACCTGCGCTTTTGGAGAAACAGCTGTGCGAGCTTTCTG 540

614 TCATGCTGTGTAACAAGAGCTTCATTTATCTGAGACAGATTAATTAAGTTTAA 673

541 TCATGCTGTGTAACAAGAGCTTCATTTATCTGAGACAGATTAATTAAGTTTAA 600

674 CTTTAAACCGCTTCACCTGCGCTTCATTTATCTGAGACAGATTAATTAAGTTTAA 733

601 CTTTAAACCGCTTCACCTGCGCTTCATTTATCTGAGACAGATTAATTAAGTTTAA 660

734 AGAAGTGAAGGAAAGCACTTCCCGACCGCGAGAGCTTTATCTGAGTGAAGGA 793

661 AGAAGTGAAGGAAAGCACTTCCCGACCGCGAGAGCTTTATCTGAGTGAAGGA 720

794 GTCTGAGTGTGATTTGCGCAGTGTGTTAACTTGTGACAGTACTGAGTGTGAGAC 853

721 GTCTGAGTGTGATTTGCGCAGTGTGTTAACTTGTGACAGTACTGAGTGTGAGAC 780

854 AAGAATGAAATGAGCTCTTCTTGAAGTGAAGAA 887

781 AAGAATGAAATGAGCTCTTCTTGAAGTGAAGAA 814

# RESULT 3

ID AA290039 standard; cDNA; 582 BP.

XX AA290039;

XX 09-MAY-2000 (first entry)

DE Hydrophobic domain containing protein clone HP02403 coding sequence.

XX Hydrophobic domain; clone HP02403; nutritional supplement; SCID; HIV;

KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;

KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;

KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;

XX anti-inflammatory; infection; bodily characteristic; ss.

XX Homo sapiens.

XX WO200000506-A2.

XX 06-JAN-2000.

XX 18-JUN-1999; 99WO-JP003242.

XX 26-JUN-1998; 98JP-00180008.

XX (SAGA) SAGAMI CHEM RES CENT.

XX (PROT-) PROTEGENE INC.

XX Kato S, Kimura T;

XX WPI: 2000-160665/14.

XX P-PSDB: AAY78802.

XX Novel human proteins having hydrophobic domains used for research and

XX diagnostic purposes.

XX Claim 3; Page 84; 117pp; English.

XX This sequence represents the hydrophobic domain containing protein, clone

CC HP00631 coding region. The sequence is isolated from a human stomach

CC cancer cell line. The HP02403 protein contains one putative transmembrane

CC domain. The protein shows homology to the Japanese quail apoptosis

CC regulator NR-13. The invention relates to human proteins with hydrophobic

CC domains, the DNA and the cDNA encoding them. The polynucleotides and

CC proteins are predicted to have biological activities which make them

CC suitable for treating, preventing or ameliorating medical conditions in

CC humans and animals. Suggested activities include nutritional activity

CC (nutritional source or supplement); cytokine and cell

CC proliferation/differentiation activity; immune stimulating (e.g. as

CC vaccines) or suppressing activity (e.g. to treat various immune

CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic

CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary

CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin

CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease

CC and autoimmune inflammatory eye disease, as well as asthma, allergies and

CC organ transplantation; haematopoiesis regulating activity (e.g. in

CC treatment of myeloid or lymphoid cell deficiencies); tissue growth

CC activity (e.g. wound healing and tissue repair, ulcers, burns,

CC periodontal disease); activating/inhibiting activity; chemotactic/chemokinetic

CC activity; haemostatic and thrombolytic activity (e.g. treating

CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and

CC tumour inhibition activity. The polynucleotides are also stated to be

CC useful for gene therapy. Other activities include inhibiting infections

CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,

CC malaria); effecting bodily characteristics such as, e.g. weight, colour,

CC skin, effecting biohythms or cardiac cycles; enhancing fertility;

CC treatment of depression; treatment of pain; hormonal or endocrine

CC activity. The polynucleotides may also be used for recombinant expression

CC of the protein

XX Sequence 582 BP; 85 A; 193 C; 191 G; 113 T; 0 U; 0 Other;

Query Match 65.6%; Score 582; DB 3; Length 582;  
Best Local Similarity 100.0%; Pred. No. 4e-121;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 80 ATGCGCAGCCGCTGCGGAGCGACCGAGCTGTTGCTGGCCGACTACCTGGAGTACTGC 139
DB 1 ATGCGCAGCCGCTGCGGAGCGACCGAGCTGTTGCTGGCCGACTACCTGGAGTACTGC 60
QY 140 GCGCGGAGACCGCGACCGCGAGCGCGCATCCAGCGCGCGCGCGCGCGCGCTGTGCGC 199
DB 61 GCGCGGAGACCGCGACCGCGAGCGCGCATCCAGCGCGCGCGCGCGCGCGCTGTGCGC 120
QY 200 TCCCGCGCGCGAGCTTACGAGATTACCGAGTCTTTTCTCCGCTTACCTCGGCTAC 259
DB 121 TCCCGCGCGCGAGCTTACGAGATTACCGAGTCTTTTCTCCGCTTACCTCGGCTAC 180
QY 260 CCGCGGAGACCGCTTACGAGTGTGCGCTGATGCGGATTCGCTCTCCGACAGCCCC 319
DB 181 CCGCGGAGACCGCTTACGAGTGTGCGCTGATGCGGATTCGCTCTCCGACAGCCCC 240
QY 320 GCGCGCAGCTGCGGAGTGTGCGCTGATGCGGATTCGCGAGGAGCGCTGTGAGAGA 379
DB 241 GCGCGCAGCTGCGGAGTGTGCGCTGATGCGGATTCGCGAGGAGCGCTGTGAGAGA 300
QY 380 GCGCGCTGTGACCGCGCGGTGAGAGAGTGGGCTTCCAGCCGCGCTAAGAGAGAG 439
DB 301 GCGCGCTGTGACCGCGCGGTGAGAGAGTGGGCTTCCAGCCGCGCTAAGAGAGAG 360
QY 440 GAGGCGAGCTGCGCGGAGATGCCAGGCGCTGTGCTGTGAGCTCGCGGCTCATG 499
DB 361 GAGGCGAGCTGCGCGGAGATGCCAGGCGCTGTGCTGTGAGCTCGCGGCTCATG 420
QY 500 GGGGAGACCGCGCTGTGAGGCTGAGGCGGCTGGAGGCTTTGTGACTTCTTC 559
DB 421 GGGGAGACCGCGCTGTGAGGCTGAGGCGGCTGGAGGCTTTGTGACTTCTTC 480
QY 560 AGACCCCGCTTTCACCTGCTTTTGGAGAAAACAGCTGTCTCAGGCTTTTCTGATGC 619
DB 481 AGACCCCGCTTTCACCTGCTTTTGGAGAAAACAGCTGTCTCAGGCTTTTCTGATGC 540
QY 620 TTGTTAACAGACGCTTATTTATCTGTGACAGCATTAATA 661
DB 541 TTGTTAACAGACGCTTATTTATCTGTGACAGCATTAATA 582
```

## RESULT 4

AAH47022 ID AAH47022 standard; cDNA; 522 BP.

AAH47022;

29-OCT-2001 (first entry)

Human Bcl-2-like polypeptide encoding cDNA (clone HLIBE40).

Bcl-2-like polypeptide; autoimmune disorder; allergy; immunomodulatory;

respiratory; cardiovascular; antirheumatic; immunostimulant; vaccine;

immunosuppressive; antiinflammatory; gene therapy; ss.

Homo sapiens.

Location/Qualifiers

Key CDS 44..448 /tag=a /product="Bcl-2-like polypeptide" /note="gene No. 2"

MO200157060-A1.

09-AUG-2001.

31-JAN-2001; 2001WO-US003080.

XX 01-FEB-2000; 2000US-0179487P.  
PR 07-FEB-2000; 2000US-0180697P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Duan DR, Ni J;  
XX WPI; 2001-476279/51.  
DR P-PSDB; AAB85666.  
XX  
XX Nucleic acids encoding human Bcl-2-like polypeptides, useful for  
PT preventing, diagnosing and/or treating.  
FT  
PS  
PS Claim 1; Page 276; 285pp; English.

CC The invention provides nucleic acid molecules (NAM1) encoding 4 human Bcl-2-like polypeptides (PEP1). The NAM1 and PEP1 may be used in the CC prevention, diagnosis and treatment of diseases associated with CC inappropriate Bcl-2-like polypeptides' expression. The NAM1 may be used CC to produce the soluble Bcl-2-like polypeptides by standard recombinant CC methodology. The polypeptides may also be used as antigens in the CC production of antibodies against Bcl-2 and in assays to identify CC modulators of Bcl-2 expression and activity. The anti-Bcl-2 antibodies CC and antagonists may be used to down regulate expression and activity. The CC anti-PEP1 antibodies may also be used as diagnostic agents for detecting CC the presence of Bcl-2 polyps in samples (e.g. by enzyme linked CC immunosorbent assay (ELISA)). Disorders that may be prevented, diagnosed CC and/or treated by the above methods include, immunodeficiencies (e.g. a CC gammaglobulinemia and B cell lymphoproliferative disorder), autoimmune CC disorders (e.g. Rheumatoid arthritis and Grave's disease), allergic CC reactions, inflammations, respiratory diseases and cardiovascular CC disorders (a full list of disorders is given in the specification). The CC present sequence represents a human Bcl-2-like polypeptide encoding cDNA

Sequence 522 BP; 102 A; 148 C; 154 G; 118 T; 0 U; 0 Other;

Query Match 57.1%; Score 506.8; DB 4; Length 522;  
Best Local Similarity 99.4%; Pred. No. 3.2e-104;  
Matches 519; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
QY 270 GCTTCAGCTGTGCGCGCTGATGCGGATTCGCTGCTCCGAGACCGCGCGCCACCT 329
DB 1 GCTTCAGCTGTGCGCGCTGATGCGGATTCGCTGCTCCGAGACCGCGCGCCACCT 60
QY 330 GGGGAGAGTGTGACGCTGCTGACCTTCGAGGAGCGCTGCTGAGAGAGGCGCTGG 389
DB 61 GGGGAGAGTGTGACGCTGCTGACCTTCGAGGAGCGCTGCTGAGAGAGGCGCTGG 120
QY 390 TGACCGCCCGGTGAGAGAGTGGGCTTCCAGCCGCGCTAAGAGAGGAGGCGAG 449
DB 121 TGACCGCCCGGTGAGAGAGTGGGCTTCCAGCCGCGCTAAGAGAGGAGGCGAG 180
QY 450 TCGCCCGGAGTGTGACGCGCTGCTGAGCTGTGAGCTGCGGCTCATGCGGAGAGC 509
DB 181 TCGCCCGGAGTGTGACGCGCTGCTGAGCTGTGAGCTGCGGCTCATGCGGAGAGC 240
QY 510 GCGCTGTGCTCAGGCTCAGGCGCGGTGAGTGTGCTTGTCACTTCTTACAGACCCCT 569
DB 241 GCGCTGTGCTCAGGCTCAGGCGCGGTGAGTGTGCTTGTCACTTCTTACAGACCCCT 300
QY 570 TTCACTGTGCTTTTGGAGAAAACAGCTGTGCAAGCTTTTCTGTATGCTTTTAACA 629
DB 301 TTCACTGTGCTTTTGGAGAAAACAGCTGTGCAAGCTTTTCTGTATGCTTTTAACA 360
QY 630 CAGCTTATTTATCTGTGACAGATTAATGAGTTTAACTTTTAACTTAACTCCGCTCT 689
DB 361 CAGCTTATTTATCTGTGACAGATTAATGAGTTTAACTTTTAACTTAACTCCGCTCT 420
QY 690 ACTGCGCACTGTGACCACTAATGACAGATGTGTGAGAACAGAACTGAGGAGAAAGC 749
DB 421 ACTGCGCACTGTGACCACTAATGACAGATGTGTGAGAACAGAACTGAGGAGAAAGC 480
```

Qy 750 ACCTTCCCCACCCGAGC-TTTTATCTGATGCATACAA 790  
|||||  
Db 481 ACCTTCCCTCACCCGAGCGTTTTTACTGATGCATACAA 522  
|||||  
RESULT 5  
ABK41913  
ID ABK41913 standard; cDNA, 548 BP.  
XX  
AC ABK41913;  
XX  
DT 21-MAY-2002 (first entry)  
XX  
DE cDNA encoding novel human connective tissue related polypeptide #301.  
XX  
KW Human; connective tissue related disorder; cancer; gene therapy;  
KM cytosolic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155343-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001332.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220966P.  
PR 26-JUL-2000; 2000US-0220966P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 16-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 06-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0231808P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234984P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249256P.

```
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCT INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX MPI; 2001-565190/63.
XX
XX P-PSDB; AAU86735.
XX
XX Nucleic acid encoding novel connective tissue associated polypeptides,
XX PT used in diagnosing, preventing, treating or ameliorating a disorder such
XX as cancer or rheumatoid arthritis.
XX
XX Claim 4; SEQ ID NO 311; 673bp; English.
XX
XX The present invention relates to the isolation of novel human connective
XX tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
XX (cDNA and genomic) sequences encoding them. The sequences of the
XX invention are useful in the diagnosis, treatment, prevention and/or
XX prognosis of diseases associated with connective tissue(s), including
XX cancer. The polynucleotide sequences of the invention are also useful in
XX gene therapy. ABR41613-ABR42101 represent cDNA sequences encoding the
XX novel human connective tissue related polypeptides. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pcr_sequences
XX
XX Sequence 548 BP; 99 A; 154 C; 167 G; 121 T; 0 U; 7 Other;
XX
XX Query Match 51.5%; Score 457.2; DB 4; Length 548;
XX Best Local Similarity 98.2%; Pred. No. 4,6e-93;
XX Matches 481; Conservative 2; Mismatches 5; Indels 2; Gaps 2;
XX
OY 259 CCCCAGGAACCGCTTCAGAGTGTGGCGCTGATGCGGAGATTCCGTCTCCGACAGCC 318
DB 61 CCACGCGTCCGCGCTTCAGAGTGTGGCGCTGATGCGGAGATTCCGTCTCCGACAGCC 120
OY 319 CGGCCCCCTGTGGGCGAAGTGTGACCTCTGTGACCTTCGAGGAGAGCTGTGGAAG 378
DB 121 CGGCCCCCACTT-GGGCAAGATGTGACCTCTGTGACCTTCGAGGAGAGCTGTGGAAG 179
OY 379 AGGCGCGTGTGACCGCCCGGTGGAAGAAGTGGGGCTTCGACGCGGCTTAAAGACA 438
DB 180 AGGCGCGCTGTGACCGCCCGGTGGAAGAAGTGGGGCTTCGACGCGGCTTAAAGACA 239
OY 439 GAGAGGCGAAGCTGCGCCGAGACTGCGACGCGCTGTGAGCTTGTGAGCTCGGCTCAT 498
DB 240 GAGAGGCGAAGCTGCGCCGAGACTGCGACGCGCTGTGAGCTTGTGAGCTCGGCTCAT 299
OY 499 GGGGAGCAACCGCGCTGTGCTCAGGCTCAGGGCGGCTGGATGTGCTTTGTCACTTCTT 558
DB 300 GGGGAGCAACCG-GCCTGTGCTCAGGCTCAGGGCGGCTGGATGTGCTTTGTCACTTCTT 358
OY 559 CAGGACCCCTTTCTCACTGGCTTTTGTGAGAAAACAGCTGTGCAAGGCTTTTCTGTCAAG 618
DB 359 CAGGACCCCTTTCTCACTGGCTTTTGTGAGAAAACAGCTGTGCAAGGCTTTTCTGTCAAG 418
OY 619 CTTGTTAAACAACGCTTCATTATCTCTGACACGATTATTATGAGTTTAAAACTTTT 678
```

```
DB 419 CTTGTTAAACAACGCTTCATTATCTCTGACACGATTATTATGAGTTTAAAACTTTT 478
OY 679 AACCGCTTCACCTGCGCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAGAAC 738
DB 479 AACCGCTTCACCTGCGCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAGAAC 538
OY 739 TGAGGGAAG 748
DB 539 TGAGGGAAG 548
XX
XX RESULT 6
XX ADB59580
XX ID ADB59580 standard; cDNA; 548 BP.
XX
XX ADB59580;
XX
XX 04-DEC-2003 (first entry)
XX
XX DE Connective tissue related polynucleotide #301.
XX
XX cyrostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;
XX antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;
XX antiinflammatory; antiallergic; antiaesthetic; dermatological;
XX nephrotoxic; virucide; fungicide; antibacterial; antiparasitic;
XX gene therapy; ds; connective tissues disorder; rheumatoid arthritis;
XX systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;
XX cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;
XX Alzheimer's disease; Parkinson's disease; cardiovascular disease;
XX atherosclerosis; myocarditis; cardiopulmonary bypass complication;
XX autoimmune disease; multiple sclerosis; allergic reaction; asthma;
XX rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;
XX gastrointestinal disorder; inflammatory bowel disease;
XX organ transplant rejection; immune system disorder; Brucella disease;
XX X-linked lymphoproliferative syndrome;
XX B-cell lymphoproliferative disorder; HIV; AIDS; infection;
XX chromosome identification; chromosome mapping;
XX connective tissue related polynucleotide; gene; ss.
XX
XX Homo sapiens.
XX
XX US2003054375-A1.
XX
XX 20-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00092154.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198125P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214866P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
```





OY	259	CCCCGGGAAACCGCTTCAGACTGCTGGCGCTGATGAGGAGATTCCTGATCTCCAGACGCC	318
Db	61	CCACGCGTCCGCGCTTCAGACTGCTGGCGCTGATGAGGATTCCTGCTCTCCAGACGCC	120
OY	319	CGGCGCCACCTTGAGGAGAGTGGTGAACGCTCGTGACCTTTCGAGGGACGCTGCTGAGAG	378
Db	121	CGGCGCCACCTTGGGAGAGTGGTGAACGCTCGTGACCTTTCGAGGGACGCTGCTGAGAG	179
OY	379	AGGGCGCGTGTGACCGCGCGGTGAGAAAGTGGGGCTTCCAGCGCGCGCTTAAGAGCA	438
Db	180	AGGCGCGCTGTGACCGCGCGGTGAGAAAGTGGGGCTTCCAGCGCGCGCTTAAGAGCA	239
OY	439	GGAGGGCGACGTGCGCCCGGGACTGCCAGCGCGCTGTGGCTTGTGAGCTCGCGGCTCAT	498
Db	240	GGAGGGCGACGTGCGCCCGGGACTGCCAGCGCGCGCTGTGGCTTGTGAGCTCGCGGCTCAT	299
OY	499	GGGGGAGACACCGCGCTGGCTGACGCTCAGGGCGGCTGGAGTGGCTTTGTGACTTCTT	558
Db	300	GGGGGAGACACCGGCTGGCTGACGCTCAGGGCGGCTGGAGTGGCTTTGTGACTTCTT	358
OY	559	CAGGACCGCCCTTTCACCTGGCTTTTGGAGAAACAGCTGGTCCAGGCTTTTCTGCATG	618
Db	359	CAGGACCGCCCTTTCACCTGGCTTTTGGAGAAACAGCTGGTCCAGGCTTTTCTGCATG	418
OY	619	CTTGTTAACAACAGCTTCACTTATATCTCTGGACACGATTAATATAGTTTAAACCTTTT	678
Db	419	CTTGTTAACAACAGCTTCACTTATATCTCTGGACACGATTAATATAGTTTAAACCTTTT	478
OY	679	AACCGGCTTCTACCTGGCCCAACTGTGACCAACTAAATGACAGATGTGTGGAACAAGAAC	738
Db	479	AACCGGCTTCTACCTGGCCCAACTGTGACCAACTAAATGACAGATGTGTGGAACAAGAAC	538
OY	739	TGAGGGAAG 748	
Db	539	TGAGGGAAG 548	

RESULT 7
ABQ44404
ID ABQ44404 standard; DNA; 874 BP.
XX
AC ABQ44404;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30995.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PE 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
XX from chemically treated DNA.
XX

PS Claim 12, 56pp + Sequence Listing; 56pp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridisation to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridised to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central  
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly by detecting mutations or single nucleotide polymorphisms  
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation  
 CC status of many C residues to be determined simultaneously. ABQ1410-  
 CC ABQ1411 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention  
 XX  
 SQ Sequence 874 BP; 144 A; 90 C; 320 G; 320 T; 0 U; 0 Other;

Query Match	36.6%;	Score 324.6;	DB 6;	Length 874;
Best Local Similarity	75.1%;	Pred. No. 3.3e-63;		
Matches 405; Conservative	0;	Mismatches 134;	Indels 0;	Gaps 0;

QY	CGGGCGAAGAAAACGAGGAAAGGCCCGGCCCCCGACGAGAGGCCGCGACATGTTGACCA	60
Db	100 CCGGTTAAGAAAATTAGCGAAGGTTCCGATTTTATAGTAGAGTCCGATTTATGTTGATT	159
QY	61 GTTGGGAGCGCACACCATGAGCCGACCGCTGCGGAGCGCACCGAGCTGTTCTGAGC	120
Db	160 GTTGGCGAGCGATTTATTATGATCGATTCGTTGCGGAGCGGATACGATGTTGTTGGT	219
QY	121 CGACTACCTGAGGGATCTGCGCCCGGGAAACCGGGACCCCCGAGCGGAGCCATTCACGCGC	180
Db	220 CGATTTATTTGGGGATTTCCGTTCCGGAAATTGGGATTTTTCGAGTCCGGCTTATTTACGTT	279
QY	181 CGAGGCGCGCGTCTGCGCTCCGCGGCCGCGCAGGTTACGGACAGATTCAACCGTCTTTT	240
Db	280 CGAGGTCTCGTGTGCGTTTCGCGCGTGGTTAGGTTACGGTATGATTTATCGGTTTTTTTT	339
QY	241 CTCGCCCTACCTCCGCTACCCCGGGAACCGGCTGAGCGTGGAGGCGATTCGCGGATTC	300
Db	340 TTTCGTTATTTCCGTTATTTTCGGGAATTCGTTTCAGATTGATGGCGTTGATGCGGATTT	399
QY	301 CGTGTCTCCGACAGCCCCCGGCCACCTGGGGCAGAGTGTGACGCTCGTACCTTCGC	360
Db	400 CGTGTTTTTCGATAGTTTTCGGTTTATTTGGGGGTGAGTGTGACGTTTCGTTATTTTCGT	459
QY	361 AGGACGCTGCTGAGAGAGGGCCGCTGTACCGCCCGGTGGAGAAAGTGGGGCTTCCA	420
Db	460 AGGACGTTGTGAGAGAGGGTCGTGTGTATCTTCGATGGAAAGATGGAGTTTATTA	519
QY	421 GCCGCGCTAAAGACGAGAGGGCGAGCTCGCCCCGGGACTGCGCAAGCGCTGTGAGCCTT	480
Db	520 GTCCGCGTTAAAGAGGTAGAGAGGCGACGTGCTTCGGGATGTGTACGTTTGTGTGTTT	579
QY	481 GCTAGCTCGCGGCTCATGATGGGCGACACCGCGCTGCTGACAGCTCACAGGCGGCTGGG	539
Db	580 GTTAGGTTTCGGGTTATGGGGTATATCCGCTTGTGTGATAGTTTAAAGCGCGGTTGGG	638

RESULT	8
ID	ABQ44405/C
XX	ABQ44405 standard; DNA, 874 BP.
AC	
XX	ABQ44405;
XX	
DT	12-JUL-2002 (first entry)

XX XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 30996.  
DE  
XX  
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
KW SNP; cell differentiation; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO200218632-A2.  
PN  
XX  
XX 07-MAR-2002.  
PD  
XX  
XX 01-SEP-2001; 2001WO-EP010074.  
PF  
XX  
XX 01-SEP-2000; 2000DE-01043826.  
PR  
XX 05-SEP-2000; 2000DE-01044543.  
PR  
XX (EPIG-) EPIGENOMICS AG.  
PA  
XX  
XX Olek A, Piepenbrock C, Berlin K, Guetig D;  
PI  
XX WPI; 2002-371829/40.  
DR  
XX  
XX Determining the degree of cytosine methylation in genomic DNA, useful for  
PT diagnosis and prognosis, comprises selective hybridization of amplicons  
PT from chemically treated DNA.  
XX  
XX Claim 12; 56pp + Sequence Listing; 56pp; German.  
XX  
XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one member,  
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
CC degree of hybridisation to both classes is determined from the label on  
CC the amplicon. From the ratio of labels hybridised to the two classes of  
CC oligomers, the degree of methylation is calculated. The method is used:  
CC (1) for diagnosis and/or prognosis of side effects of therapeutic drugs  
CC and of a wide range of diseases, e.g. cancer, disorders of the central  
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
CC particularly by detecting mutations or single nucleotide polymorphisms  
CC (SNP's); and (11) for differentiation of cell or tissue types and for  
CC investigating cell differentiation. The method allows the methylation  
CC status of many C residues to be determined simultaneously. ABQ13410-  
CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
CC for determining the degree of cytosine methylation described in the  
CC disclosure of the invention  
XX  
XX Sequence 874 BP; 320 A; 320 C; 90 G; 144 T; 0 U; 0 Other;  
SQ  
Query Match 36.6%; Score 324.6; DB 6; Length 874;  
Best Local Similarity 75.1%; Pred. No. 3,3e+63;  
Matches 405; Conservative 0; Mismatches 134; Indels 0; Gaps 0;  
QY 1 CGGGCCAAAGAAACAGGAGGCGGCGGCCCCCGAGAGAGCGGACCATGTTGACCA 60  
DB 775 CGGGTTAAGAAATATAGCGAAGTTTCGGTTTTTATGTAAGGTCCGATTATGTTGATTA 716  
QY 61 GTTCCGGAGAGGACACCATAGCCGCGCTGGGGAGGACCGAGCTTTGCTGGC 120  
DB 715 GTTGGGGAGCGTATATATATATGTCGATTCTGGGAGAGCGTATCGAGTTGTTCGT 656  
QY 121 CGACTACTGGGGTACTGCGCCCGGGAACCGGACCGCCCGGAGCGGAGCCATCCAGCC 180  
DB 655 CGATTATTTGGGGTATTCGCTTCGGGAATTCGATATTTTCAGTGGCGGTTATTTACGT 596  
QY 181 CGAGGCGCGCTGCTGCGCTCCGCGGCGGCGGAGTTACGCGAGATTCAACCGTCTCTTTT 240  
DB 595 CGAGGTCGTGCTGTTGCGTTTCGCGGTGTTAGTTAGTTACGTTATTCGTTTTCCTT 536

QY 241 CTCGCTTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGGCGCTGATGCGGATTC 300  
DB 535 TTTCGTTATTTTCGTTATTTTCGGAATCTGTTTCAGTTGTTGGTGTGATGCGGATTT 476  
QY 301 CGTGCTTCGACAGACCCCGGCCCCCAGCTGGGAGAGTGTGACGCTCGTACCTTCGC 360  
DB 475 CGTGTTCGATAGTTTCGTTTATTTGGGGGTGAGTGTGACGTTTCGTTATTTTCGT 416  
QY 361 AGGACGCTGCTGAGAGAGAGGCGCTGTACCCCGGTGGAGAGTGGGCTTCCA 420  
DB 415 AGGACGTTGTGAGAGAGGCGCTGTGTATGTTTCGTTGAGAGAGTGGGCTTTTA 356  
QY 421 GCCGCGCTAAAGAGAGAGGCGAGCTGCGCCCGGAGCTGCCAGCGCTGTGGCCTT 480  
DB 355 GTCCGCGTTAAAGAGTGAAGGCGGAGCGTCTGTTCCGATTTGTTAGCCTTTGGTGT 296  
QY 481 GCTGAGCTCGCGGCTCATGAGGAGAGACCGGCGCTGCTGACGCTCAGGCGGCTGGG 539  
DB 295 GTTGAATTTCGCGGTTTATGAGGTTATGCGGTTTGTGTTGTTAGGTTAGGCGGTTTGG 237  
RESULT 9  
ABQ44403 standard; DNA; 874 BP.  
ID ABQ44403;  
XX  
XX ABQ44403;  
AC  
XX  
XX 12-JUL-2002 (first entry)  
DT  
XX  
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 30994.  
DE  
XX  
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
KW SNP; cell differentiation; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO200218632-A2.  
PN  
XX  
XX 07-MAR-2002.  
PD  
XX  
XX 01-SEP-2001; 2001WO-EP010074.  
PF  
XX  
XX 01-SEP-2000; 2000DE-01043826.  
PR  
XX 05-SEP-2000; 2000DE-01044543.  
PR  
XX (EPIG-) EPIGENOMICS AG.  
PA  
XX  
XX Olek A, Piepenbrock C, Berlin K, Guetig D;  
PI  
XX WPI; 2002-371829/40.  
DR  
XX  
XX Determining the degree of cytosine methylation in genomic DNA, useful for  
PT diagnosis and prognosis, comprises selective hybridization of amplicons  
PT from chemically treated DNA.  
XX  
XX Claim 12; 56pp + Sequence Listing; 56pp; German.  
XX  
XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one member,  
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
CC degree of hybridisation to both classes is determined from the label on  
CC the amplicon. From the ratio of labels hybridised to the two classes of  
CC oligomers, the degree of methylation is calculated. The method is used:  
CC (1) for diagnosis and/or prognosis of side effects of therapeutic drugs  
CC and of a wide range of diseases, e.g. cancer, disorders of the central  
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,



Db 355 ACCGGGACTAAAAAAGACAGCTGCGCCGAAATACCAACGCTTATACCTT 296  
Qy 481 GCTGAGCTCGGCGCTCATGGGACAGCAGCGCTGTGCTGACAGCTCAGGCGCT 536  
Cc 295 ACTAACTCGGACTCATTAACACACCGCGCTTACTTACCAACTCAAAAGACT 240  
Db  
RESULT 11  
ADD34132  
ID ADD34132 standard; DNA; 650 BP.  
AC ADD34132;  
XX 15-JAN-2004 (first entry)  
DT  
XX Mouse mitochondrial DNA sequence SEQ ID NO:1910.  
DE  
XX ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
KM mitochondrial disease; oxidative phosphorylation dysfunction;  
KW oxidative stress; apoptosis; aging.  
XX  
OS Mus musculus.  
XX MO2003020220-A2.  
XX 13-MAR-2003.  
PD 30-AUG-2002; 2002MO-US027886.  
PF 30-AUG-2001; 2001US-0316323P.  
PR 31-AUG-2001; 2001CA-02356540.  
XX  
PA (UDEM-) UNIV EMORY.  
PI Wallace DC, Levy S, Kerstam K, Procaccio V;  
DR WPI; 2003-300821/29.  
XX  
PT Array containing probes for genes involved in mitochondrial biology,  
PT useful for determining mitochondrial biology gene expression profiles for  
PT use in diagnosing pathologies and identifying biochemical pathways.  
XX  
PS Claim 2; SEQ ID NO 1910; 201pp; English.  
XX  
CC The invention relates to a novel array comprising at least two isolated  
CC nucleotide molecules, each molecule having a sequence capable of uniquely  
CC hybridising to a nucleic acid molecule which is an expression product of  
CC a gene involved in mitochondrial biology. The array comprises two or more  
CC isolated nucleic acid molecules or spots, each molecule having a sequence  
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array  
CC of the invention is useful for determining an expression profile of a  
CC mouse or human sample containing nucleic acid, by contacting the array  
CC with the sample under conditions allowing selective hybridisation, and  
CC measuring hybridisation of nucleic acid in the sample to the array to  
CC produce an expression profile. The array is also useful for determining  
CC an expression profile of a first labelled sample containing nucleic acid  
CC relative to a second, differently labelled sample containing nucleic  
CC acid. The second sample is a reference or a standard. An array is useful  
CC for determining an expression profile diagnostic of an energy-metabolism-  
CC related physiological condition. An array of the invention is useful for  
CC determining mitochondrial biology gene expression profiles of organisms,  
CC such as human, mice and closely related species, tissue and organs of  
CC such organisms, which are useful for determining expression profiles  
CC diagnostic of energy metabolism-related physiological conditions,  
CC diagnosing such physiological conditions, identifying biochemical  
CC pathways, genes, and mutations involved in such physiological conditions,  
CC identifying therapeutic agents useful for preventing and/or treating such  
CC physiological conditions, evaluating and/or monitoring the efficacy of  
CC such therapies, and creating and identifying animal models of human  
CC energy metabolism-related physiological conditions. An array is also  
CC useful for defining expression signatures or profiles for mitochondrial  
CC diseases, as well as distinguishing clinical disorders that result from  
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,

CC apoptosis and aging. An array of the invention contains probes of genes  
CC not previously recognised to participate in mitochondrial biology. The  
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA  
CC clones used to make the probes of the invention. Some sequences are not  
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,  
CC 1906, 2408 and 2643.  
XX  
SQ Sequence 650 BP; 143 A; 172 C; 170 G; 165 T; 0 U; 0 Other;  
SQ  
Query Match 27.1%; Score 240.2; DB 10; Length 650;  
Best Local Similarity 64.7%; Pred. No. 3e-44; Indels 25; Gaps 3;  
Matches 417; Conservative 0; Mismatches 203;  
Qy 86 GACCGCTGCGGAGCCGACCGAGCTTGTGCGCCGACTACCTGAGGCTACGCGCCG 145  
Cc 17 GACCCACTGACATGACAGCAGCAGCAGCGCTGCTGTCTATCATATTTCTTGGCGCAAG 76  
Db 146 GAACCGGACCCCGGAGCCGCGCCATCAAGCCCGAGGCGCGCTGCGCTCGCG 205  
Qy 77 GAGCGGACACCCGAGACCCGCGCCAGCTGTGCGAGGCGCGCTTGTGCTGTG 136  
Db 206 GCCGCAAGTTACGCAATTCACCGGCTTTTCTCCGCTACCTGCGCTAACCCGCG 265  
Qy 137 ACTAGGACGATTCAGCAGAGCAGCAGCAGATTTTCTCTCTTCTGAAAGCCGCGGC 196  
Db 266 AACCGCTTGAGCTGTGCGCTGATGAGCGGATTCGCTCTCGACAGCCCGCGCC 325  
Qy 197 AATGCCCTGAGCTGTGTAAACAGATGCAATAGTTGCTCTCCAAAGACCAAGACTTC 256  
Db 326 ACTGCGGAGAGATGATGAGCTGTGCTGCTTGTGAGGAGCGCTGAGAGAGGCGCG 385  
Qy 257 AGCTGAGACCAATGATGATCTCTGCGGAGGAGCGCTTATGATTAACAGGCGCT 316  
Db 386 CTGATGACCGCGCGGTGAAGAAGTGGGCTTCCAGCGCGCTTAAAGAGAGAGGCG 445  
Qy 317 TACATGCTGTCTAGCAAGAA-----GGATCTGGAGATCTGTTC 358  
Db 446 GACGTCGCCGAGACTGCGCAGCGCTGTGCTGTGCTGCTGCGCTCATGAGG--- 502  
Qy 359 ATAATGACCCAGAGCTGCTGTCTATGATGAATTTCTGTATATCTGCTCATGAGGCGT 418  
Db 503 CAGCAGCGCGCTGCTGAGGCTCAAGGCGCGCTGAGATGCTTTGTCACTTCTTCAGG 562  
Qy 419 CGGACCGCGCAGGCTGAGGCTCTGCGCGCTGAGATGCTTTGTCCGCTTCTTCAG 478  
Db 563 ACCCTTTTCACTGCTTTTGTGAGAAACAGCTGTGCGGCTTTTGTGATGCTTG 622  
Qy 479 AATCTTTTACCGCTGCTGAGGATTTGATTCAGGCTTTTGTGATGAGGCTTC 538  
Db 623 TTAACAAGCGCTTATTTATCTGAGCAGATTTATGATTTTAAACCTTTAAC 682  
Qy 539 TTGCAACAGCGCTTTTATCTGGAAGAGTTAT---AAGTTTAAATTTTAAAG 594  
Db 683 CGCTTCTACCTGCGCACTGAGCACTAATGATGATGATGATG 727  
Qy 595 CGATTTTACCTTACCTGATGAGCCCTTAAGGCAATTCGG 639  
Db  
RESULT 12  
ADD34582/C  
ID ADD34582 standard; DNA; 835 BP.  
AC ADD34582;  
XX 15-JAN-2004 (first entry)  
DT  
XX Mouse mitochondrial DNA sequence SEQ ID NO:2360.  
DE  
XX ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
KM mitochondrial disease; oxidative phosphorylation dysfunction;  
KW oxidative stress; apoptosis; aging.  
XX  
OS Mus musculus.

XX WO2003020220-A2.  
 PN 13-MAR-2003.  
 PD 30-AUG-2002; 2002WO-US027886.  
 PF 30-AUG-2001; 2001US-0316323P.  
 PR 31-AUG-2001; 2001CA-02356540.  
 XX (UYEM-) UNIV EMORY.  
 PA Wallace DC, Levy S, Kerstam K, Procaccio V;  
 PI WPI; 2003-300821/29.  
 DR WPI; 2003-300821/29.  
 XX  
 XX Array containing probes for genes involved in mitochondrial biology,  
 PT useful for determining mitochondrial biology gene expression profiles for  
 PT use in diagnosing pathologies and identifying biochemical pathways.  
 XX  
 PS Claim 2; SEQ ID NO 2360; 201pp; English.

CC The invention relates to a novel array comprising at least two isolated  
 CC nucleotide molecules, each molecule having a sequence capable of uniquely  
 CC hybridising to a nucleic acid molecule which is an expression product of  
 CC a gene involved in mitochondrial biology. The array comprises two or more  
 CC isolated nucleic acid molecules or spots, each molecule having a sequence  
 CC chosen from sequence of 994 human probes and 2046 mouse probes. An array  
 CC of the invention is useful for determining an expression profile of a  
 CC mouse or human sample containing nucleic acid, by contacting the array  
 CC with the sample under conditions allowing selective hybridisation, and  
 CC measuring hybridisation of nucleic acid in the sample to the array to  
 CC produce an expression profile. The array is also useful for determining  
 CC an expression profile of a first labelled sample containing nucleic acid  
 CC relative to a second, differently labelled sample containing nucleic  
 CC acid. The second sample is a reference or a standard. An array is useful  
 CC for determining an expression profile diagnostic of an energy-metabolism-  
 CC related physiological condition. An array of the invention is useful for  
 CC determining mitochondrial biology gene expression profiles of organisms,  
 CC such as human, mice and closely related species, tissue and organs of  
 CC such organisms, which are useful for determining expression profiles  
 CC diagnostic of energy metabolism-related physiological conditions,  
 CC diagnosing such physiological conditions, identifying biochemical  
 CC pathways, genes, and mutations involved in such physiological conditions,  
 CC identifying therapeutic agents useful for preventing and/or treating such  
 CC physiological conditions, evaluating and/or monitoring the efficacy of  
 CC such therapies, and creating and identifying animal models of human  
 CC energy metabolism-related physiological conditions. An array is also  
 CC useful for defining expression signatures or profiles for mitochondrial  
 CC diseases, as well as distinguishing clinical disorders that result from  
 CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,  
 CC apoptosis and aging. An array of the invention contains probes of genes  
 CC not previously recognised to participate in mitochondrial biology. The  
 CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA  
 CC clones used to make the probes of the invention. Some sequences are not  
 CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,  
 CC 1906, 2408 and 2643.  
 XX  
 XX Sequence 835 BP; 238 A; 200 C; 185 G; 212 T; 0 U; 0 Other;

Query Match 17.4%; Score 154; DB 10; Length 835;  
 Best Local Similarity 62.7%; Pred. No. 7; 7e-25;  
 Matches 291; Conservative 0; Mismatches 165; Indels 8; Gaps 3;  
 QY 378 GAGGGCCGCTGTCACCCCGGTGAAGAATGGGGCTTCCAGCCGCGCTTAAGAGAC 437  
 DB 816 GCGGACACTTATGATCAAGGCCCTTACATGCTGTCAAGAGAAAGGAGATCTGGGGA 757  
 QY 438 AGGAGGGGAGCTCGCCGCGGAGCTGCAGCGGCTGTGGCTGAGCTCGGGCTCA 497  
 DB 756 ATCTGTATATGATGACCCGAGACTGTCTCTATATGATGACATTTCTGTAATATCTGCTCA 697  
 QY 498 TGGGG---CAGCACCGCGCTGTGCAAGGCTCAAGGGCGGCTGGGATGGCTTTTGTCACT 554

DB 696 TGGGGCGTCGACCGGCGCAGCTGAGGCTCTCGGGGCTGGGATGGCTTTGGCCGCT 637  
 QY 555 TCTTACAGACCCCTTTTCCATCTGCTTTTGAAGAAAACACTGCTCCAGGCTTTTCTGT 614  
 DB 636 TCTTCAAAATCTCTTACCGCTCGGCTGTGAGAAAGTTCTGATTCAGGCTTTTCTGT 577  
 QY 615 CATGCTTGTAAACAACAGCTTTCATTTATCTGTGACACGATTTATGAGTTTAAAC 674  
 DB 576 CAGGCTTTTTCGACAGCCATCTTTTATCTGGAACGTTTAT---AAGTTTAAAT 521  
 QY 675 TTTTAAACCGCTTCTACCTGCGCAACTGTGACCACTAATATGACAGATGTGTGAGAACAA 734  
 DB 520 TTTTAAAGCGATCTTACCTACCTACCTGTGAACCTCTAAGAAACATTCGGGTAAATG 461  
 QY 735 GAACGTAGAGAAAGACCTTCCCAACCCAGAGCTTTTATCTGTAATGATCAAGAG 794  
 DB 460 GATTAAGAACTGAGAAAGTCCCTCAGCCTTAGAGACATTTCTACCTGCAATGC-TACATGAG 402  
 QY 795 TCTGAGGTGATGATTTGGCCAGTGTTTTAACTGTGACAAGTA 838  
 DB 401 TCTGAGTTTGTGATGGGCGCAGCATTTGAAGGGGCCCATGTA 358

RESULT 13  
 ADD34583  
 ID ADD34583 standard; DNA; 338 BP.  
 AC ADD34583;  
 XX  
 XX 15-JAN-2004 (first entry)  
 DT  
 XX  
 XX Mouse mitochondrial DNA sequence SEQ ID NO:2361.

XX ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
 KW mitochondrial disease; oxidative phosphorylation dysfunction;  
 KW oxidative stress; apoptosis; aging.  
 OS Mus musculus.

PN WO2003020220-A2.  
 XX  
 XX 13-MAR-2003.  
 PD 30-AUG-2002; 2002WO-US027886.  
 PF 30-AUG-2001; 2001US-0316323P.  
 PR 31-AUG-2001; 2001CA-02356540.  
 XX  
 XX (UYEM-) UNIV EMORY.

PI Wallace DC, Levy S, Kerstam K, Procaccio V;  
 DR WPI; 2003-300821/29.  
 XX  
 XX WPI; 2003-300821/29.

PT Array containing probes for genes involved in mitochondrial biology,  
 PT useful for determining mitochondrial biology gene expression profiles for  
 PT use in diagnosing pathologies and identifying biochemical pathways.  
 XX  
 XX Claim 2; SEQ ID NO 2361; 201pp; English.

CC The invention relates to a novel array comprising at least two isolated  
 CC nucleotide molecules, each molecule having a sequence capable of uniquely  
 CC hybridising to a nucleic acid molecule which is an expression product of  
 CC a gene involved in mitochondrial biology. The array comprises two or more  
 CC isolated nucleic acid molecules or spots, each molecule having a sequence  
 CC chosen from sequence of 994 human probes and 2046 mouse probes. An array  
 CC of the invention is useful for determining an expression profile of a  
 CC mouse or human sample containing nucleic acid, by contacting the array  
 CC with the sample under conditions allowing selective hybridisation, and  
 CC measuring hybridisation of nucleic acid in the sample to the array to  
 CC produce an expression profile. The array is also useful for determining  
 CC an expression profile of a first labelled sample containing nucleic acid

CC relative to a second, differently labelled sample containing nucleic  
CC acid. The second sample is a reference or a standard. An array is useful  
CC for determining an expression profile diagnostic of an energy-metabolism-  
CC related physiological condition. An array of the invention is useful for  
CC determining mitochondrial biology gene expression profiles of organisms,  
CC such as human, mice and closely related species, tissue and organs of  
CC such organisms, which are useful for determining expression profiles  
CC diagnostic of energy metabolism-related physiological conditions,  
CC diagnosing such physiological conditions, identifying biochemical  
CC pathways, genes, and mutations involved in such physiological conditions,  
CC identifying therapeutic agents useful for preventing and/or treating such  
CC physiological conditions, evaluating and/or monitoring the efficacy of  
CC such therapies, and creating and identifying animal models of human  
CC energy metabolism-related physiological conditions. An array is also  
CC useful for defining expression signatures or profiles for mitochondrial  
CC diseases, as well as distinguishing clinical disorders that result from  
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,  
CC apoptosis and aging. An array of the invention contains probes of genes  
CC not previously recognized to participate in mitochondrial biology. The  
CC sequences shown in A003324-A0035260 represent murine mitochondrial DNA  
CC clones used to make the probes of the invention. Some sequences are not  
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,  
CC 1906, 2408 and 2643.

XX Sequence 338 BP; 76 A; 95 C; 100 G; 67 T; 0 U; 0 Other;

Query Match 10.0%; Score 88.6; DB 10; Length 338;

Best Local Similarity 64.3%; Pred. No. 3.2e-10;

Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 86 GACCCGCTGCGGAGAGCGAGCTGTGTGCGCGGAGCTAGTCCGCCCG 145  
Db 132 GACCCAGTCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 191  
Qy 146 GAACCCGAGCCCGGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 205  
Db 192 GAGCGGAG 251  
Qy 206 GCGCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 265  
Db 252 ACTAGGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 311  
Qy 266 AACCGCTGAGTGTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 292  
Db 312 AATGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338

RESULT 14

ADA71938  
ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant, bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.

XX Oryza sativa.

XX MO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001MO-IB001105.

XX 22-JUN-2001; 2001MO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI Kategiri F, Quan S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;  
XX WPI: 2003-175290/17.

DR Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 7.7%; Score 68; DB 8; Length 2000;

Best Local Similarity 11.0%; Pred. No. 2e-05;

Matches 86; Conservative 362; Mismatches 322; Indels 12; Gaps 3;

Qy 7 AAGAAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 66  
Db 26 MMSCAFMSSHMSRMSRMSRMSRMSRMSRMSRMSRMSRMSRMSRMSR 85  
Qy 67 GAG 126  
Db 86 SGKMKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR 145  
Qy 127 CCGGAG 186  
Db 146 MWCARAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 199  
Qy 187 GCGGAG 246  
Db 200 CGRSGCGRSAYSRYSRYGTSRYGTSRYGTSRYGTSRYGTSRYGTSRYG 259  
Qy 247 CTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306  
Db 260 SMMRMKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR 319  
Qy 307 CTCGAG 366  
Db 320 SRMAMTKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR 379  
Qy 367 GCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426  
Db 380 SRTSRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR 439  
Qy 427 GCTAAG 486  
Db 440 GMMRMKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR 499  
Qy 487 CTCGAG 544  
Db 500 CRRRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR 559  
Qy 545 TTTTGTCACTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604  
Db 560 KRMYYKMMWYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 619  
Qy 605 ----GCTTTCTGCACTTTGTTAACAAGAGAGAGAGAGAGAGAGAG 660  
Db 620 CRKASRRSSAKRYAMGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679  
Qy 661 ATGAGTTTAAACTTTAAACCCGCTTCTACCTGCAAGAGAGAGAGAGAG 720





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2005, 10:37:53 ; Search time 60 seconds

(without alignments)  
1741.069 Million cell updates/sec

Title: US-10-071-174a-2

Perfect score: 1084  
Sequence: 1 MVDQLRETTMADPLERTE.....QAFSLCLTTFATYIMTRL 204

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1035	95.5	194	1	BCLB_HUMAN
2	487	44.9	185	1	O99M66
3	474.5	43.8	191	1	BCLB_MOUSE
4	467.5	43.1	191	2	Q7TPY8
5	217.5	20.1	179	2	O9B1F2
6	193.5	17.9	177	1	NR13_COTUA
7	188.5	17.4	177	2	O90ZNI
8	178.5	16.5	162	2	O9DH00
9	176.5	16.3	176	2	O8UWDS
10	174.5	16.1	176	2	O8UWDS
11	143.5	13.2	211	2	O9W6F1
12	131	12.1	192	2	O8HYU5
13	126	11.6	192	2	O8S043
14	125	11.5	174	2	O9W6F2
15	123	11.3	173	2	O90KJ3
16	122	11.3	350	2	BAXA_BOVIN
17	122	11.3	350	2	O7YRZ9
18	121.5	11.2	188	2	O6GP82
19	121	11.2	192	1	BAXA_RAT
20	120	11.1	164	2	O70AD3
21	120	11.1	192	1	BAXA_HUMAN
22	118	10.9	239	2	O7SSV7
23	117	10.8	173	2	O8K3J2
24	117	10.8	192	1	BAXA_MOUSE
25	117	10.8	193	1	BCLM_MOUSE
26	117	10.8	229	1	BCL2_BOVIN
27	116	10.7	173	2	O8WZ49
28	116	10.7	235	2	O81008
29	115	10.6	192	2	O6A093
30	115	10.6	193	1	BCLM_MOUSE
31	114.5	10.6	238	2	O90Z98

32	113	10.4	114	2	O9NR76	O9NR76 homo sapien
33	113	10.4	143	1	BAXD_HUMAN	P55269 homo sapien
34	113	10.4	178	2	O8CFR2	O8CFR2 mus musculu
35	113	10.4	178	2	O9CYW5	O9CYW5 mus musculu
36	112.5	10.4	228	1	ARI_XENLA	O91827 xenopus lae
37	111.5	10.3	175	2	O925A9	O925A9 rattus norv
38	111	10.2	193	2	O88996	O88996 rattus norv
39	111	10.2	219	2	O7TS60	O7TS60 rattus norv
40	111	10.2	350	2	O8HY55	O8HY55 canis famli
41	109	10.1	239	1	BCL2_HUMAN	P10415 homo sapien
42	108	10.0	228	2	O8UWDS	O8UWDS brachydantio
43	107.5	9.9	192	2	O919N4	O919N4 brachydantio
44	107	9.9	193	2	O8CG14	O8CG14 mus musculu
45	107	9.9	350	1	MCCL1_HUMAN	O07820 homo sapien

## ALIGNMENTS

RESULT 1  
BCLB\_HUMAN  
ID BCLB\_HUMAN STANDARD: PRT: 194 AA.  
AC O9HD36; Q8TCS9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUN-2004 (Rel. 44, Last annotation update)  
DE Apoptosis regulator Bcl-2 (Bcl-2-like 10 protein) (Anti-apoptotic protein Nth).  
GN Name=BCL2L10; Synonym=BCLB;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=21548034; PubMed=11689480; DOI=10.1093/hmg/10.21.2329;  
RA Zhang H., Holgrave W., De Geyter C.;  
RT "Bcl-2, a novel anti-apoptotic member of the Bcl-2 family, blocks apoptosis in the mitochondria death pathway but not in the death receptor pathway.";  
RT Hum. Mol. Genet. 10:2329-2339(2001).  
RL (2)  
RN SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=21201065; PubMed=11278245; DOI=10.1074/jbc.C000871200;  
RA Ke N., Godzik A., Reed J.C.;  
RT "Bcl-2, a novel Bcl-2 family member that differentially binds and regulates Bax and Bak.";  
RL J. Biol. Chem. 276:12481-12484(2001).  
RN (3)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21477277; PubMed=11593390; DOI=10.1038/sj/onc.1204740;  
RA Aouachria A., Arnaud E., Venet S., Lalle P., Gony M., Rigal D., Gillet G.;  
RT "Nth, a human homologue of Nr-13 associates with Bcl-2s and is an inhibitor of apoptosis.";  
RL Oncogene 20:5846-5855(2001).  
RN (4)  
RP FUNCTION: Promotes cell survival. Suppresses apoptosis induced by BAX but not BAK.  
RN (5)  
RP SUBUNIT: Binds to Bcl-2, Bcl-X and BAX. Interacts with APAF-1.  
RN (6)  
RP SUBCELLULAR LOCATION: Associated with mitochondria and the nuclear envelope.  
RN (7)  
RP TISSUE SPECIFICITY: Widely expressed in adult tissues.  
RN (8)  
RP Preferentially expressed in the lungs, the liver and the kidneys.  
RN (9)  
RP SIMILARITY: Belongs to the Bcl-2 family.  
RN (10)  
RP SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.  
RN (11)  
RP SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.  
RN (12)  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF102501; AAD08703.1; -  
 CC EMBL; AF067660; AAC8150.1; -  
 DR MGD; MGI:1330841; Bcl2l11.0.  
 DR GO; GO:0006916; P:anti-apoptosis; IDA.  
 DR GO; GO:0006915; P:apoptosis; IDA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PS50062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH1\_FALSE\_NEG.  
 DR PROSITE; PS01258; BH2\_FALSE\_NEG.  
 KM Apoptosis; Mitochondrion; Transmembrane.  
 FT DOMAIN 79 98 BH1.  
 FT DOMAIN 144 155 BH2.  
 FT TRANSMEM 166 183 Potential.  
 SQ SEQUENCE 191 AA; 22302 MW; 819014E6B2DFE411 CRC64;

Query Match 43.8%; Score 474.5; DB 1; Length 191;  
 Best Local Similarity 48.2%; Pred. No. 2.3e-36;  
 Matches 94; Conservative 32; Mismatches 56; Indels 13; Gaps 3;

QY 13 DPLRRETELLADYICGAREGTEPEAPSTPEAAVLNSAARLQIHRSFSAVLGYPG 72  
 DB 6 DPLHRTLRLLSDYIFPCAREPDTPPEPTSVREALLSVTRQIQOEHQEFSSCSERG 65  
 QY 73 NRELVALMADSVSDSPGPTWGRVITLTPFAGTLLEGPVLTARMKKGPGPRLKEQEG 132  
 DB 66 NRELVKMGADKXLSKQDFSWQLVMLAFAGTLMGPGVPAVAKK-----DLG 116  
 QY 133 D---VARDCCRLVALLSSRLMG-QHRAMLQAGGWDGCHFRPTPEPLAFMRKOLVOAFL 188  
 DB 117 NRVIYTRCCCLIVNLYMLMGRRRRARLEALGWDGCFRPFKNPLPLGFWRGLLIQVFL 176  
 QY 189 SCLLTAFIYLMTRL 203  
 DB 177 SGFPATALFTWKRL 191

## RESULT 4

Q7TPY8 PRELIMINARY; PRT; 191 AA.  
 AC Q7TPY8;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Bcl2l10 protein.  
 DE Mus musculus (Mouse).  
 OS Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Egg;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J., Boeck S.A., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W., Villalón D.K., Muzny D.M., Sodergren B.U., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues Y., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S., Krzywinski M.I., Skalska U., Smalios D.B., Scherch A., Schein J.E., RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Egg;  
 RA Strausberg R.;  
 RX Submitted (May-2003) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; BC052690; AAH52690.1; -  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR GO; GO:0006916; P:anti-apoptosis; IDA.  
 DR GO; GO:0006915; P:apoptosis; IDA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PS50062; BCL2\_FAMILY; 1.  
 SQ SEQUENCE 191 AA; 22230 MW; F2F176E6B0DE69A6 CRC64;

Query Match 43.1%; Score 467.5; DB 2; Length 191;  
 Best Local Similarity 47.7%; Pred. No. 1e-35;  
 Matches 93; Conservative 32; Mismatches 57; Indels 13; Gaps 3;

QY 13 DPLRRETELLADYICGAREGTEPEAPSTPEAAVLNSAARLQIHRSFSAVLGYPG 72  
 DB 6 DPLHRTLRLLSDYIFPCAREPDTPPEPTSVREALLSVTRQIQOEHQEFSSCSERG 65  
 QY 73 NRELVALMADSVSDSPGPTWGRVITLTPFAGTLLEGPVLTARMKKGPGPRLKEQEG 132  
 DB 66 NRELVKMGADKXLSKQDFSWQLVMLAFAGTLMGPGVPAVAKK-----DLG 116  
 QY 133 D---VARDCCRLVALLSSRLMG-QHRAMLQAGGWDGCHFRPTPEPLAFMRKOLVOAFL 188  
 DB 117 NRVIYTRCCCLIVNLYMLMGRRRRARLEALGWDGCFRPFKNPLPLGFWRGLLIQVFL 176  
 QY 189 SCLLTAFIYLMTRL 203  
 DB 177 SGFPATALFTWKRL 191

## RESULT 5

Q9ELP2 PRELIMINARY; PRT; 179 AA.  
 AC Q9ELP2;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE RSOEFL N13-like protein.  
 GN Name=RSORFL;  
 OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 CC Alphaherpesvirinae; Mardivirus.  
 NCBI\_TaxID=37108;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FC126;  
 RX MEDLINE=21195611; PubMed=11297687;  
 RA Kingham B.F., Zelnik V., Kopacek J., Majerciak V., Ney E., Schmidt C.J.;  
 RT "The genome of herpesvirus of turkeys: comparative analysis with  
 RT Marek's disease viruses."  
 RT J. Gen. Virol. 82:1123-1135 (2001).  
 DR EMBL; AF282130; AAG30102.1; -  
 DR HSSP; 007817; 1R21.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR Pfam; PF00452; BCL-2; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PS50062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01258; BH2; UNKNOWN\_1.

SQ SEQUENCE 179 AA; 19011 MW; A872B038A96D1823 CRC64;  
 Query Match 20.1%; Score 217.5; DB 2; Length 179;  
 Best Local Similarity 35.7%; Pred. No. 2.6e-12;  
 Matches 61; Conservative 14; Mismatches 55; Indels 41; Gaps 6;  
 QY 11 MADPLRERTELLADYLGVCAREPPTPEAPSTPEAAVLRSAAARLQIHRSFF--SAVL 68  
 DB 1 MADSLKETALLLEDFYQHCCKEG---PPPS-PTAAELRRAAALERRERPFRRSCAPL 56  
 QY 69 GYGNRFEFLVAM--ADSVLSDSPGPTGWRVTVLTPAGTL-----LERGPLVTARWK 119  
 DB 57 ASGGTQALISALQSIVSELSNGSGFRWGRCLATIVLGGSLATATLYENGCEBGP----- 109  
 QY 120 KMGFPRLKEGEGDVARDQRLVALLSSRLMGCHRAVLQAGGWDGFCRPF 170  
 DB 110 -----SRLAALATAYLAEBEGEWMEHGGMDGFCRPF 141

RESULT 6  
 NR13 COTUA STANDARD; PRT; 177 AA.  
 AC Q90343;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Apoptosis regulator NR-13.  
 GN Name=NR-13;  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 OX NCBI\_TaxID=93934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Fibroblast, and Neuroretina;  
 RX MEDLINE=95246730; PubMed=7729415;  
 RA Gillet G., Guerin M., Trembleau A., Brun G.;  
 RT "A Bcl-2-related gene is activated in avian cells transformed by the  
 RT Rous sarcoma virus".  
 RL EMBL J. 14:1372-1381(1995).  
 CC -1- FUNCTION: Rous sarcoma virus-activated protein with anti-apoptotic  
 CC properties.  
 CC -1- SUBCELLULAR LOCATION: Plasma membrane.  
 CC -1- TISSUE SPECIFICITY: Mainly expressed in neural and muscular  
 CC tissues.  
 CC -1- DEVELOPMENTAL STAGE: Expression is dramatically down-regulated  
 CC after embryonic day 7 in the optic tectum, and correlates with the  
 CC onset of apoptosis in this area.  
 CC -1- SIMILARITY: Belongs to the Bcl-2 family.  
 CC -1- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.  
 CC -1- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X84418; CA459136.1; -.  
 DR HSSP; 007817; IMAZ.  
 DR InterPro; IPR000712; BCL2\_BH.  
 DR InterPro; IPR002475; BCL2\_Family.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR PROSITE; PS50062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH1; 1.  
 DR PROSITE; PS01258; BH2; 1.  
 KW Apoptosis; Transmembrane.  
 FT DOMAIN 75 94 BH1.  
 FT DOMAIN 126 141 BH2.  
 FT TRANSMEM 86 106 Potential.

FT TRANSMEM 156 177 Potential.  
 SQ SEQUENCE 177 AA; 18826 MW; ADSCE79D3353C1F CRC64;  
 Query Match 17.9%; Score 193.5; DB 1; Length 177;  
 Best Local Similarity 33.1%; Pred. No. 4.5e-10;  
 Matches 57; Conservative 16; Mismatches 56; Indels 43; Gaps 6;  
 QY 11 MADPLRERTELLADYLGVCAREPPTPEAPSTPEAAVLRSAAARLQIHRSFFSAVL-- 68  
 DB 1 MGSLSKETALLLEDFYQHCCKEG---PPPS-PTAAELRRAAALERRERPFRRSCAPL 56  
 QY 69 --GYGNRFEFLVAM--ADSVLSDSPGPTGWRVTVLTPAGTL-----LERGPLVTARW 118  
 DB 57 ARAEPRBAALLRKVAALQLETDG--GLNWGRLLATLVFAGTLAALAESACEBGP----- 109  
 QY 119 KMGFPRLKEGEGDVARDQRLVALLSSRLMGCHRAVLQAGGWDGFCRPF 170  
 DB 110 -----SRLAALATAYLAEBEGEWMEHGGMDGFCRPF 141

RESULT 7  
 Q90ZNI PRELIMINARY; PRT; 177 AA.  
 AC Q90ZNI;  
 DT 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Anti-apoptotic NR13.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22304624; PubMed=12133006; DOI=10.1042/BJ20020836;  
 RX Lalle P., Aouacheria A., Dumont-Miscopein A., Jamon M., Venet S.,  
 RA Bobichon H., Colas P., Deleage G., Geoufjon C., Gillet G.;  
 RT "Evidence for crucial electrostatic interactions between Bcl-2  
 RT homology domains BH3 and BH4 in the anti-apoptotic Nr-13 protein.";  
 RL Biochem. J. 368:213-221(2002).  
 DR EMBL; AF375661; AAK54806.1; -.  
 DR HSSP; 007817; IMAZ.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR SMART; SM00337; BCL2\_FAMILY; 1.  
 DR PROSITE; PS50062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01258; BH2; 1.  
 SQ SEQUENCE 177 AA; 18778 MW; F06EP9C8B1AF250D CRC64;  
 Query Match 17.4%; Score 188.5; DB 2; Length 177;  
 Best Local Similarity 34.9%; Pred. No. 1.3e-09;  
 Matches 60; Conservative 12; Mismatches 57; Indels 43; Gaps 7;  
 QY 11 MADPLRERTELLADYLGVCAREPPTPEAPSTPEAAVLRSAAARLQIHRSFF--SAVL 68  
 DB 1 MGSLSKETALLLEDFYQHCCKEG---PPPS-PTAAELRRAAALERRERPFRRSCAPL 56  
 QY 69 GYGNRFEFLVAM--ADSVLSDSPGPTGWRVTVLTPAGTL-----LERGPLVTARW 118  
 DB 57 ARAEPRBAALLRKVAALQLETDG--GLNWGRLLATLVFAGTLAALAESACEBGP----- 109  
 QY 119 KMGFPRLKEGEGDVARDQRLVALLSSRLMGCHRAVLQAGGWDGFCRPF 170  
 DB 110 -----SRLAALATAYLAEBEGEWMEHGGMDGFCRPF 141

RESULT 8  
 Q9DH00 PRELIMINARY; PRT; 162 AA.  
 AC Q9DH00;  
 DT 01-MAR-2001 (TRENBLrel. 16, Created)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Similar to Apoptosis inhibitor, Bcl2 family proteins (Apoptosis  
 inhibitor-like protein).  
 GN Name=HT079; Synonym=HT096;  
 OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Mardivirus.  
 CX NCBI\_TaxID=37108;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FC126;  
 RX MEDLINE=2057832; PubMed=1134310; DOI=10.1128/JVI.75.2.971-978.2001;  
 RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.F.;  
 RT "The genome of turkey herpesvirus."  
 RL J. Virol. 75:971-978(2001).  
 DR EMBL; AF291866; AAG45828.1; -.  
 DR EMBL; AF291866; AAG45819.1; -.  
 DR HSSP; Q07817; IR21.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2 BH.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PSS0062; BCL2\_FAMILY; 1.  
 SQ SEQUENCE 162 AA; 17286 MW; 97CE08A7834EA58B CRC64;

Query Match 16.5%; Score 178.5; DB 2; Length 162;  
 Best Local Similarity 33.5%; Pred. No. 1e-08;  
 Matches 55; Conservative 14; Mismatches 54; Indels 41; Gaps 6;

QY 11 MADPLRETELLADYLGICAREPPTPEAPSTPEAAVLRSAARLRQIHRSPF--SYVL 68  
 DB 1 MADSLKETALLLEHYPOHCCKEG---PPPS-PRALERRAAALERERFRFFSCAPL 56  
 QY 69 GYPGNRFELVALM--DSVLSDSPGPTWGRVTVLTVPAGTL-----LERGPLVTARW 119  
 DB 57 ASGGTQALSLQSVSELSGSGFNWGRCLATIVLGSLATVLYENGCEBGP----- 109  
 QY 120 KMGQFQPLKEQBGDVARDCCRLVALLSRLMGQHRAMLQAGGMDGFCRPF 163  
 DB 110 -----SRLAALAAVLAEGEWLEHARWDGFCRPF 134

RESULT 9

08UWD5 PRELIMINARY; PRT; 176 AA.  
 AC 08UWD5;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Nr13.  
 GN Name=mc11; Synonym=nr13;  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 CX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2037392; PubMed=10917738;  
 RA Inohara N., Nunez G.;  
 RT "Genes with homology to mammalian apoptosis regulators identified in  
 zebrafish."  
 RL Cell Death Differ. 7:509-510(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Inohara N., Nunez G.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF41285; AAL32471.1; -.  
 DR HSSP; Q07817; IR21.  
 DR ZFIN; ZDB-GENE-030825-2; mc11.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR Pfam; PF00452; Bcl-2; 1.

DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PSS0062; BCL2\_FAMILY; 1.  
 DR PROSITE; PSS0258; BH2; UNKNOWN 1.  
 SQ SEQUENCE 176 AA; 19677 MW; 26CDE124AE77F573 CRC64;

Query Match 16.3%; Score 176.5; DB 2; Length 176;  
 Best Local Similarity 31.8%; Pred. No. 1.7e-08;  
 Matches 55; Conservative 21; Mismatches 46; Indels 51; Gaps 9;

QY 15 LRETELLADYLGICAREPPTPEAPSTPEAAVLRSAARLRQIHRSPFSA---YLCY 70  
 DB 5 LREQTLLAEADYISFCSGIQQTP--PSES---AENARRYLAKEMEQHRKTRSLQSFLDT 60  
 QY 71 PGNRFELVALMAD-----SVLSDSPGPTWGRVTVLTVPAGT-----LERGPLVTARW 117  
 DB 61 CG-----ADPSKCIQSVARELVGDKRWGVSIFFTTGYLASBLSRG----- 105  
 QY 118 KMGQFQPLKEQBGDVARDCCRLVALLSRLMGQHRAMLQAGGMDGFCRPF 170  
 DB 106 -----ENSEGS-----SRLAETIADYLGCEKQDMVLVENGWEGFCRPF 143

RESULT 10

0804D0 PRELIMINARY; PRT; 177 AA.  
 AC 0804D0;  
 DT 01-JUN-2003 (TReMBLrel. 24, Created)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Bcl-2 related anti-apoptotic protein.  
 GN Name=Nr13;  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 CX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Neiman P.E.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY186728; AAO31808.1; -.  
 DR HSSP; Q07817; IMAZ.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2 BH.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR Pfam; PF00452; BCL-2; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PSS0062; BCL2\_FAMILY; 1.  
 SQ SEQUENCE 177 AA; 18891 MW; B94BP9CF0A8C282 CRC64;

Query Match 16.1%; Score 174.5; DB 2; Length 177;  
 Best Local Similarity 33.7%; Pred. No. 2.7e-08;  
 Matches 58; Conservative 12; Mismatches 59; Indels 43; Gaps 7;

QY 11 MADPLRETELLADYLGICAREPPTPEAPSTPEAAVLRSAARLRQIHRSPF--SYVL 68  
 DB 1 MPGLKETALLLEHYPOH--RAGGALPPSAT--AAELRRRAAALERERFRFFSCAPL 56  
 QY 69 GYPGNRFELVALM--ADSVLSDSPGPTWGRVTVLTVPAGT-----LERGPLVTARW 118  
 DB 57 ARAEPR-EAALLLRVAQAQLEAGGLNGRLLALVVFPTGLAALAESECEBGP----- 109  
 QY 119 KMGQFQPLKEQBGDVARDCCRLVALLSRLMGQHRAMLQAGGMDGFCRPF 170  
 DB 110 -----SRLAALAAVLAEGEWLEHARWDGFCRPF 141

RESULT 11

Q9W6F1 PRELIMINARY; PRT; 211 AA.  
 AC Q9W6F1;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)

DT 01-MAR-2004 (Tremblrel. 26, last annotation update)  
 DE Myeloid cell leukemia protein MCL-1 (Fragment).  
 OS Gallus gallus (Chicken).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=90311;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99190706; PubMed=10090728;  
 RA Lee R.W., Gillet G., Burnside J., Thomas S.J., Neinan P.,  
 RT "Role of Nr13 in regulation of programmed cell death in the bursa of  
 RL Fabricsius.";  
 RN Genes Dev. 13:718-728 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Soter L., Burnside J.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF120210; AAD31644.1; -;  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH; 1.  
 DR PROSITE; PS01258; BH2; UNKNOWN\_1.  
 FT NON\_TER  
 SQ SEQUENCE 211 AA; 23143 MW; A4B05763F92BAC7 CRC64;

Query Match 13.2%; Score 143.5; DB 2; Length 211;  
 Best Local Similarity 28.5%; Pred. No. 2.7e-05;  
 Matches 53; Conservative 21; Mismatches 61; Indels 51; Gaps 9;

QY 13 DPLRRTRELLADYGYCARF-----PCTPEPAPSTPEAAVRSAAAR 57  
 DB 16 DELRQESLELLRLRYREAGAEAPGVKLPGLGGPGRGRASS---AVWEKLETLR 71  
 QY 58 QIHRFPFSAF-LGYG-----NRELVALMADSVLSDSPGPT-WGRVVTLTVPAG 105  
 DB 72 RVGDGVQGHBLAFQGMRLKRLIKKEDLIQAVCEVAQVFNQ--GVTNMGRVVTLTISF-- 127  
 QY 106 TLLERGPVLTAAWKWGFQPRLKEQGVARDQORLVALLSRLMGORHRAWLQAQGWMDG 165  
 DB 128 -----GAFPAKKHLK-----INQKCTIT---SLAGITTLAVSSKKEWLMSQGWEG 171  
 QY 166 FCHFR 171  
 DB 172 FVDFFR 177

RESULT 12  
 Q8HYU5 PRELIMINARY; PRT; 192 AA.  
 AC Q8HYU5;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)  
 DE Bax.  
 OS Canis familiaris (Dog).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yazawa M., Masuda K., Ohno K., Tsujimoto H.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB080230; BACS3619.1; -;  
 DR HSP; 007812; 1F16.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR Pfam; PF00452; Bcl-2; 1.

DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH; 1.  
 DR PROSITE; PS01258; BH2; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 SQ SEQUENCE 192 AA; 21299 MW; 7B3C326DDE62A87E CRC64;

Query Match 12.1%; Score 131; DB 2; Length 192;  
 Best Local Similarity 25.8%; Pred. No. 0.00035;  
 Matches 56; Conservative 29; Mismatches 70; Indels 62; Gaps 11;

QY 10 TWADPLRRTRELLADYGYCARFPG--TPP-PAPSTPEAAVRSAAARLQIHRFSFA 66  
 DB 14 TSSQIMTKGALLLQGFIDRAGRGRGGETPELPEQVPODASTKLSCLKRIQIDELS- 72  
 QY 67 YLGYGNRFELVALMADSVLSDSPG-----TWGRVVTLTVPAGTLLER 111  
 DB 73 -----NMELQRMIA-AVDTSPREVFRRVALEMFSQGNFMGRVVALFFPASKLVKA 124  
 QY 112 -----PLVTAARKWKFQPRLKEQGVARDQORLVALLSRLMGORHRAWLQAQGWMDG 166  
 DB 125 LCTKVPBELRTIMGW-----TLDFLRRLTG-----WIQDQGWMDL 161  
 QY 167 CHFFRTPEPLAFWKQVLAFLSLITTAFTYLTWL 203  
 DB 162 LSYFGTP-----TW-QVTITFVAGVL-TASILTWKGM 191

## RESULT 13

Q8SQ43 PRELIMINARY; PRT; 192 AA.

AC Q8SQ43;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, last annotation update)  
 DE Bax-protein.  
 GN Name=bax;  
 OS Felis silvestris catus (Cat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamazaki J., Ogura K., Kano R., Hasegawa A.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB080724; BAB85810.1; -;  
 DR HSP; 007812; 1F16.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR002475; BCL2\_FAMILY.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH; 1.  
 DR PROSITE; PS01258; BH2; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 SQ SEQUENCE 192 AA; 21283 MW; 852D271AB86923FB CRC64;

Query Match 11.6%; Score 126; DB 2; Length 192;

Best Local Similarity 25.8%; Pred. No. 0.001; 70; Indels 62; Gaps 11;

Matches 56; Conservative 29; Mismatches 70; Indels 62; Gaps 11;

QY 10 TWADPLRRTRELLADYGYCARFPG--TPP-PAPSTPEAAVRSAAARLQIHRFSFA 66  
 DB 14 TSSQIMTKGALLLQGFIDRAGRGRGGETPELPEQVPODASTKLSCLKRIQIDELS- 72  
 QY 67 YLGYGNRFELVALMADSVLSDSPG-----TWGRVVTLTVPAGTLLER 111  
 DB 73 -----NMELQRMIA-AVDTSPREVFRRVALEMFSQGNFMGRVVALFFPASKLVKA 124  
 QY 112 -----PLVTAARKWKFQPRLKEQGVARDQORLVALLSRLMGORHRAWLQAQGWMDG 166  
 DB 125 LCTKVPBELRTIMGW-----TLDFLRRLTG-----WIQDQGWMDL 161



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 6, 2005, 10:39:24 ; Search time 42 Seconds

(without alignments)  
362.581 Million cell updates/sec

Title: US-10-071-174a-2

Perfect score: 1084  
Sequence: 1 MVDLRETTMADPLRERTE.....QAFSLCLTTFYIWTLL 204

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	11.2	192	1	US-08-112-208C-3 Sequence 3, Appl1
2	121	11.2	192	1	US-08-248-819A-3 Sequence 3, Appl1
3	121	11.2	192	2	US-08-337-646A-3 Sequence 3, Appl1
4	121	11.2	192	2	US-08-856-531-3 Sequence 3, Appl1
5	121	11.2	192	2	US-08-856-034-3 Sequence 3, Appl1
6	121	11.2	192	3	US-08-927-326-3 Sequence 3, Appl1
7	121	11.2	192	4	US-09-379-820A-3 Sequence 3, Appl1
8	120	11.1	192	4	US-08-607-269-25 Sequence 25, Appl1
9	120	11.1	192	1	US-08-471-058-13 Sequence 13, Appl1
10	120	11.1	192	2	US-08-856-531-9 Sequence 9, Appl1
11	120	11.1	192	2	US-08-856-034-9 Sequence 9, Appl1
12	120	11.1	192	3	US-08-471-057-13 Sequence 13, Appl1
13	120	11.1	192	3	US-09-127-048-7 Sequence 7, Appl1
14	120	11.1	192	4	US-08-470-865-13 Sequence 13, Appl1
15	120	11.1	192	4	US-09-155-327G-14 Sequence 14, Appl1
16	120	11.1	192	5	PCT-US95-04600-25 Sequence 25, Appl1
17	120	11.1	331	4	US-09-033-525-2 Sequence 2, Appl1
18	119	11.0	192	1	US-08-112-208C-9 Sequence 9, Appl1
19	119	11.0	192	1	US-08-248-819A-9 Sequence 9, Appl1
20	119	11.0	192	2	US-08-337-646A-9 Sequence 9, Appl1
21	119	11.0	192	4	US-08-927-326-9 Sequence 9, Appl1
22	119	11.0	192	4	US-09-379-820A-9 Sequence 9, Appl1
23	118	10.9	192	1	US-08-112-208C-2 Sequence 2, Appl1
24	118	10.9	192	1	US-08-248-819A-2 Sequence 2, Appl1
25	118	10.9	192	2	US-08-337-646A-2 Sequence 2, Appl1
26	118	10.9	192	2	US-08-856-531-2 Sequence 2, Appl1
27	118	10.9	192	2	US-08-856-034-2 Sequence 2, Appl1

28	118	10.9	192	3	US-08-927-326-2 Sequence 2, Appl1
29	118	10.9	192	4	US-09-379-820A-2 Sequence 2, Appl1
30	117	10.8	192	1	US-08-112-208C-8 Sequence 8, Appl1
31	117	10.8	192	1	US-08-248-819A-8 Sequence 8, Appl1
32	117	10.8	192	2	US-08-337-646A-8 Sequence 8, Appl1
33	117	10.8	192	2	US-08-856-531-8 Sequence 8, Appl1
34	117	10.8	192	3	US-08-856-034-8 Sequence 8, Appl1
35	117	10.8	192	3	US-09-127-048-6 Sequence 6, Appl1
36	117	10.8	192	3	US-08-927-326-8 Sequence 8, Appl1
37	117	10.8	192	4	US-09-379-820A-8 Sequence 8, Appl1
38	117	10.8	193	4	US-09-155-327G-7 Sequence 7, Appl1
39	117	10.8	193	4	US-09-949-016-10928 Sequence 10928, A
40	115	10.6	193	4	US-09-155-327G-9 Sequence 9, Appl1
41	115	10.6	333	4	US-09-155-327G-10 Sequence 10, Appl1
42	115	10.6	365	4	US-09-149-476-696 Sequence 696, App
43	115	10.6	365	4	US-09-010-147B-24 Sequence 24, Appl1
44	112	10.3	192	1	US-08-798-897-6 Sequence 6, Appl1
45	112	10.3	192	2	US-08-978-523-6 Sequence 6, Appl1

## ALIGNMENTS

```
RESULT 1
US-08-112-208C-3
; Sequence 3, Application US/08112208C
; Patent No. 5691179
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, Stanley J.
; TITLE OF INVENTION: CELL DEATH REGULATORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/112,208C
; FILING DATE: 26-AUG-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; LENGTH: 192 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; NAME/KEY: Protein
; LOCATION: 1..192
; OTHER INFORMATION: /note= "Protein sequence of murine
; ; OTHER INFORMATION: Bax."
US-08-112-208C-3

Query Match 11.2%; Score 121; DB 1; Length 192;
Best Local Similarity 26.7%; Pred. No. 5.8e-06;
Matches 54; Conservative 22; Mismatches 64; Indels 62; Gaps 11;

QY 23 LADYGYCARBGPPE-PAPSTPEAAVLSAARLQIHRSFSAVLYGPNRRFLVALM 81
```

Db 31 IODRAGMAGE--TELELEQPPODASTKXSECLRIIDELDS-----NMELQMI 80  
QY 82 ADVSLSDSGP-----TWGRVVTLTVPAGTLLERG-----PLVTARWKX 121  
Db 81 AD-VDRSDPREVFVAVADMFDGFMNGRVVAFYFASKVLKXCTKVELRTTIMGW 139  
QY 122 GFQPLKQEGDVARDCORLVALSSRLMGHRAWLQAGGMDGCFHFRTPPLAFWRK 181  
Db 140 -----TIDFLERLLG---WIDQGGWEGLLSTFGTP---TW-- 170  
QY 182 QLVQAFSLCLTTAFIYLMTRL 203  
Db 171 QVTITFVAGVL-TASLTITWKX 191

RESULT 2  
US-08-248-819A-3  
; Sequence 3, Application US/08248819A  
; Patent No. 5700638  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew.  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/248, 819A  
; FILING DATE: 25-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/112,208  
; FILING DATE: 26-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000610  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..192  
; OTHER INFORMATION: /note= "Protein sequence of murine  
; OTHER INFORMATION: Bax."  
US-08-248-819A-3

Query Match 11.2%; Score 121; DB 1; Length 192;  
Best Local Similarity 26.7%; Pred. No. 5.8e-06;  
Matches 54; Conservative 22; Mismatches 64; Indels 62; Gaps 11;

QY 23 LADVLGYCARBGTPE-PAPSTPEAAVLSAARLQIHRSEFSAVLGYGNGRFFELVALM 81  
Db 31 IODRAGMAGE--TELELEQPPODASTKXSECLRIIDELDS-----NMELQMI 80  
QY 82 ADVSLSDSGP-----TWGRVVTLTVPAGTLLERG-----PLVTARWKX 121

Db 81 AD-VDRSDPREVFVAVADMFDGFMNGRVVAFYFASKVLKXCTKVELRTTIMGW 139  
QY 122 GFQPLKQEGDVARDCORLVALSSRLMGHRAWLQAGGMDGCFHFRTPPLAFWRK 181  
Db 140 -----TIDFLERLLG---WIDQGGWEGLLSTFGTP---TW-- 170  
QY 182 QLVQAFSLCLTTAFIYLMTRL 203  
Db 171 QVTITFVAGVL-TASLTITWKX 191

RESULT 3  
US-08-337-646A-3  
; Sequence 3, Application US/08337646A  
; Patent No. 5856171  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/337, 646A  
; FILING DATE: 10-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/248, 819  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/112,208  
; FILING DATE: 26-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000620  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..192  
; OTHER INFORMATION: /note= "Protein sequence of murine  
; OTHER INFORMATION: Bax."  
US-08-337-646A-3

Query Match 11.2%; Score 121; DB 2; Length 192;  
Best Local Similarity 26.7%; Pred. No. 5.8e-06;  
Matches 54; Conservative 22; Mismatches 64; Indels 62; Gaps 11;

QY 23 LADVLGYCARBGTPE-PAPSTPEAAVLSAARLQIHRSEFSAVLGYGNGRFFELVALM 81  
Db 31 IODRAGMAGE--TELELEQPPODASTKXSECLRIIDELDS-----NMELQMI 80  
QY 82 ADVSLSDSGP-----TWGRVVTLTVPAGTLLERG-----PLVTARWKX 121

Db 81 AD-VDROSPREVFPRVAADMFADGNFNGRVVALLFYFASKVLKALCTKPELIRITMGW 139  
Qy 122 GFQRLKEQBGDVARDQRLVALLSSRLMGHRAMLQAGSGDGCFFRTPPPLAFWRK 181  
Db 140 -----TLDFLRERLLG-----WIDQSGWEGLLSYFGTP-----TW-- 170  
Qy 182 QLVQAFSLCLTTAFIYLMTRL 203  
Db 171 QTVITIFVAGVL-TASLTIWKKM 191

## RESULT 4

US-08-856-531-3  
; Sequence 3, Application US/08856531  
; Patent No. 5942490  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howell & Haferkamp, L.C.  
; STREET: 7733 Foreyth Blvd., Suite 1400  
; CITY: St. Louis  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/856,531  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, Donald R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 976176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..192  
; OTHER INFORMATION: /note= "Murine BAX polypeptide"  
; US-08-856-531-3

Query Match 11.2%; Score 121; DB 2; Length 192;  
Best Local Similarity 26.7%; Pred. No. 5.8e-06;  
Matches 54; Conservative 22; Mismatches 64; Indels 62; Gaps 11;  
Qy 23 LADYLYGCAREPGPE-PAPSTPEAAVIRSAARLRQIHRSFSAIYLYGPNRPELVALM 81  
Db 31 IQDRAGRWAGE--TPBELTLBQPPQDASTKKLSICLRIGDELDS-----NMELQPMI 80  
Qy 82 ADVSLDSPGPG-----TWGRVTVLTFAGTLLEBG-----PLVYARWKW 121  
Db 81 AD-VDROSPREVFPRVAADMFADGNFNGRVVALLFYFASKVLKALCTKPELIRITMGW 139  
Qy 122 GFQRLKEQBGDVARDQRLVALLSSRLMGHRAMLQAGSGDGCFFRTPPPLAFWRK 181  
Db 140 -----TLDFLRERLLG-----WIDQSGWEGLLSYFGTP-----TW-- 170  
Qy 182 QLVQAFSLCLTTAFIYLMTRL 203  
Db 171 QTVITIFVAGVL-TASLTIWKKM 191

Db 171 QTVITIFVAGVL-TASLTIWKKM 191

## RESULT 5

US-08-856-034-3  
; Sequence 3, Application US/08856034  
; Patent No. 5955595  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howell & Haferkamp, L.C.  
; STREET: 7733 Foreyth Blvd., Suite 1400  
; CITY: St. Louis  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/856,034  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, Donald R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 976175  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..192  
; OTHER INFORMATION: /note= "Murine BAX polypeptide"  
; US-08-856-034-3

Query Match 11.2%; Score 121; DB 2; Length 192;  
Best Local Similarity 26.7%; Pred. No. 5.8e-06;  
Matches 54; Conservative 22; Mismatches 64; Indels 62; Gaps 11;

Qy 23 LADYLYGCAREPGPE-PAPSTPEAAVIRSAARLRQIHRSFSAIYLYGPNRPELVALM 81  
Db 31 IQDRAGRWAGE--TPBELTLBQPPQDASTKKLSICLRIGDELDS-----NMELQPMI 80  
Qy 82 ADVSLDSPGPG-----TWGRVTVLTFAGTLLEBG-----PLVYARWKW 121  
Db 81 AD-VDROSPREVFPRVAADMFADGNFNGRVVALLFYFASKVLKALCTKPELIRITMGW 139  
Qy 122 GFQRLKEQBGDVARDQRLVALLSSRLMGHRAMLQAGSGDGCFFRTPPPLAFWRK 181  
Db 140 -----TLDFLRERLLG-----WIDQSGWEGLLSYFGTP-----TW-- 170  
Qy 182 QLVQAFSLCLTTAFIYLMTRL 203  
Db 171 QTVITIFVAGVL-TASLTIWKKM 191

## RESULT 6

US-08-927-326-3  
; Sequence 3, Application US/08927326  
; Patent No. 6184202  
; GENERAL INFORMATION:

```

GENERAL INFORMATION:
APPLICANT: Korsmeyer, Stanley J.
TITLE OF INVENTION: Cell Death Regulator
FILE REFERENCE: 6029-1314
CURRENT APPLICATION NUMBER: US/09/379, 820A
CURRENT FILING DATE: 1999-08-24
PRIOR APPLICATION NUMBER: 08/856, 034
PRIOR FILING DATE: 1997-05-14
NUMBER OF SEQ ID NOS: 32
SEQ ID NO: 3
LENGTH: 192
TYPE: PRT
ORGANISM: Mus musculus
US-09-379-820A-3

Query Match          11.2%; Score 121; DB 4; Length 192;
Best local Similarity 26.7%; Pred. No. 5,8e-06;
Matches 54; Conservative 22; Mismatches 64; Indels 62; Gaps 11

QY      23 LADYIGCYCARREGCTPE-PAPSTPEAAVLRSAAARLROIHSSFFSAYLYGPGNPELVALM 81
Db      31 IORARARMNGE--TPELTLEOEPPODASTKLSSECIKRTIGDELD-----NMELQRM 80

QY      82 ADSVLSDSGP-----TWGRVVTLVTFAGTLLERG-----PLYTARWKW 121
Db      81 AD-VDSDSREVEFFRVADMFADGNFMNGRVVAFYFASKLVKALCTKVPBELIRTMGW 139

QY      122 GFOPRLKEGEVDVARDCORLVALLSSRLMGCHRAVLQAGWGDFCHFFFTPPPLAFWRK 161
Db      140 -----TLDFFLERILG-----WIDOGGMEGLSLFYGRP-----TW-- 170

QY      182 QLVQARLSCLLTAPAFVLMTRL 203
Db      171 QVTVTFVAGVL-TASLTWKW 191

RESULT 8
US-08-607-269-25
Sequence 25, Application US/08607269
Patent No. 5702897
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Sato, Takaaki
TITLE OF INVENTION: Interaction of Proteins Involved in a
TITLE OF INVENTION: Cell Death Pathway
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/607,269
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/226,876
FILING DATE: 13-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9882
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 25:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-607-269-25

```

Query Match	11.1%;	Score 120;	DB 1;	Length 192;
Best Local Similarity	25.3%;	Pred. No. 7.7e-06;		
Matches 55;	Conservative 29;	Mismatches 71;	Indels 62;	Gaps 11

```

QY      67 YLGPNGNRELVALMADSVLSDSPGB-----TWGRVYTVTVTFACTLLERG 111
      |||
      |||
      |||
Db      73 -----NMELQMIIA-AVDTDSPREVFPRVAADMFSDONFWMGVVALPYFAASKVLVKA 124
      |||
      |||
      |||
QY      112 -----PLVTARKMKMGFQPRLKEQESDVARDCCRLVALLSRHMGRAHMTQAQSGMDGF 166
      :
      :
      :
Db      125 LCTKVVELIRITMGW-----TLDFLRERLIG---WTQDGGWDGL 161
      :
      :
      :
QY      167 CHFERTPFPLAFWRKQLQAFLSCLLTTAFYILMTRL 203
      :
      :
      :
Db      162 LSYFGTP---TW--QTVTTFVAGVLT-TASLTITMKOM 191
      :
      :
      :

```

RESULT 9  
US-08-471-058-13

GENERAL INFORMATION:  
ADDITANT, vinfofor Michael C

1	TITLE OF INVENTION:	NOVEL APOPTOSIS MODULATING
2	TITLE OF INVENTION:	PROTEINS, DNA ENCODING THE PROTEINS AND METHODS OF USE
3	TITLE OF INVENTION:	THEREOF
4	NUMBER OF SEQUENCES:	24

ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastEO for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: 08/320,157  
FILING DATE: 07-OCT-1994  
APPLICATION NUMBER: 08/160,067  
FILING DATE: 30-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lehman, Susan K  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 23647-20007,123  
TELECOMMUNICATION INFORMATION:

```

; INFORMATION FOR SEQ ID NO: 13
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 192 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;

```

US-08-471-058-13

Query Match	11.1%;	Score 120;	DB 1;	Length 192;
Best Local Similarity	-25.3%;	Pred. No. 7.7e-06;		
Matches	55;	Conservative	29;	Mismatches 71;
				Indels 62;
				Gaps 11;

[illegible]

RESULT 10  
US-08-856-531-9  
; Sequence 9, Application US/08856531

APPLICANT: KORSMEYER, Stanley J.

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

ADDRESSEE: Howell & Haterkamp, L.C.  
STREET: 7733 Forsyth Blvd., Suite 1400  
CITY: St. Louis

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: Patentin Release #1.0, Version #1.30
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/856,531

```

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, Donald R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 97  
TELECOMMUNICATION INFORMATION:

```

? INFORMATION FOR SEQ ID NO: 9:
? SEQUENCE CHARACTERISTICS:
?     LENGTH: 192 amino acids
?     TYPE: amino acid
?     STRANDEDNESS: single
?     TOPOLOGY: linear
?     MOLECULE TYPE: protein
?     FEATURE:
?         NAME/KEY: Protein
?         LOCATION: 1..192
?     OTHER INFORMATION: /note="Human Bax polypeptide"
?
?
? JS-08-856-531-9

```

	Query Match	11.1%	Score 120	DB 2	Length 192
	Best Local Similarity	25.3%	Pred. No. 7.7e-06		
	Matches 55	Conservative 29	Mismatches 71	Indels 62	Gaps 11
QY	10	TMADPLAERTELLADYIGYCARPG--TPBPA-PSTPEBAVLSAAARLQIHRSFSA	66		
	:::	:::	:::	:::	:::

Db 14 TSSEQIMTKTALLQGFIDRAGRMGGAPELALDPVPODASTKKLSCLKRIIGBELDS- 72  
QY 67 YLGGNRPFLVALMADSVLSDSPGP-----TWGRVVTLVTFAGTLLBERG 111  
Db 73 -----NMELORMIA-AVDITDSPREVFRVAADMFSDGNFMWGRVVALFYFASKLVYKA 124  
QY 112 -----PLVTARWKKGFPRLKEQBGDVARDQRLVALLSRLMGCHRAMIQAQGGWDF 166  
Db 125 LCTKVPBELRTIMGN-----TLDFLRERLLG-----WIQDQGWDGL 161  
QY 167 CHERTPPPLAFWRKQLVOAFISCLTTAFIYLMTRL 203  
Db 162 LSYFGTP-----TW--QVTITFVAGVL-TASLTIWKKM 191

RESULT 11  
US-08-856-034-9  
; Sequence 9, Application US/08856034  
; Patent No. 595595  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: CELL DEATH REGULATORS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howell & Haferkamp, L.C.  
; STREET: 7733 Foreyth Blvd., Suite 1400  
; CITY: St. Louis  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/856,034  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, Donald R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 976175  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..192  
; OTHER INFORMATION: /note= "Human BAX polypeptide"  
US-08-856-034-9

Query Match 11.1%; Score 120; DB 2; Length 192;  
Best Local Similarity 25.3%; Pred. No. 7.7e-06;  
Matches 55; Conservative 29; Mismatches 71; Indels 62; Gaps 11;

QY 10 TMADPLRRTTELLADYLGVCAREPG--TPRPA-PSTPEAAVLRSAARLRQIHRSFSA 66  
Db 14 TSSEQIMTKTALLQGFIDRAGRMGGAPELALDPVPODASTKKLSCLKRIIGBELDS- 72  
QY 67 YLGGNRPFLVALMADSVLSDSPGP-----TWGRVVTLVTFAGTLLBERG 111  
Db 73 -----NMELORMIA-AVDITDSPREVFRVAADMFSDGNFMWGRVVALFYFASKLVYKA 124  
QY 112 -----PLVTARWKKGFPRLKEQBGDVARDQRLVALLSRLMGCHRAMIQAQGGWDF 166

Db 125 LCTKVPBELRTIMGN-----TLDFLRERLLG-----WIQDQGWDGL 161  
QY 167 CHERTPPPLAFWRKQLVOAFISCLTTAFIYLMTRL 203  
Db 162 LSYFGTP-----TW--QVTITFVAGVL-TASLTIWKKM 191

RESULT 12  
US-08-471-057-13  
; Sequence 13, Application US/08471057  
; Patent No. 6015687  
; GENERAL INFORMATION:  
; APPLICANT: KIEFER, MICHAEL C.  
; APPLICANT: BARR, PHILIP J.  
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
; TITLE OF INVENTION: ENCODING THE PROTEINS AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,057  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/320,157  
; FILING DATE: 07-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LEHNHARDT, SUSAN K.  
; REGISTRATION NUMBER: 33,943  
; REFERENCE/DOCKET NUMBER: 23647-20007, 20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600  
; TELEFAX: (415) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-471-057-13

Query Match 11.1%; Score 120; DB 3; Length 192;  
Best Local Similarity 25.3%; Pred. No. 7.7e-06;  
Matches 55; Conservative 29; Mismatches 71; Indels 62; Gaps 11;

QY 10 TMADPLRRTTELLADYLGVCAREPG--TPRPA-PSTPEAAVLRSAARLRQIHRSFSA 66  
Db 14 TSSEQIMTKTALLQGFIDRAGRMGGAPELALDPVPODASTKKLSCLKRIIGBELDS- 72  
QY 67 YLGGNRPFLVALMADSVLSDSPGP-----TWGRVVTLVTFAGTLLBERG 111  
Db 73 -----NMELORMIA-AVDITDSPREVFRVAADMFSDGNFMWGRVVALFYFASKLVYKA 124  
QY 112 -----PLVTARWKKGFPRLKEQBGDVARDQRLVALLSRLMGCHRAMIQAQGGWDF 166  
Db 125 LCTKVPBELRTIMGN-----TLDFLRERLLG-----WIQDQGWDGL 161  
QY 167 CHERTPPPLAFWRKQLVOAFISCLTTAFIYLMTRL 203  
Db 162 LSYFGTP-----TW--QVTITFVAGVL-TASLTIWKKM 191

```

; RESULT 13
; US-09-127-048-7
; Sequence 7, Application US/09127048
; Patent No. 6165732
; GENERAL INFORMATION:
; APPLICANT: Koremeyer, Stanley J.
; APPLICANT: Schlesinger, Paul H.
; TITLE OR INVENTION: Method for identifying Apoptosis Modulating Compounds
; FILE REFERENCE: 6029-6052
; CURRENT APPLICATION NUMBER: US/09/127,048
; CURRENT FILING DATE: 1998-07-31
; EARLIER APPLICATION NUMBER: 60/061,823
; EARLIER FILING DATE: 1997-10-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-127-048-7

Query Match      11.1%; Score 120; DB 3; Length 192;
Beat Local Similarity 25.3%; Pred. No. 7, 7e-06;
Matches 55; Conservative 29; Mismatches 71; Indels 62; Gaps 11

Cy      10 TMADPLRRETELLADYIGYCARREG--TPRPA-ESTPEAAVLRSAAARLRQIHRSFSPA 66
      14 TSSEQIMRTGALLLOGFIODRAGRNGGEAPFELADPVQDASTKLSBCKRIIDELDS- 72
      67 YLGYGNKPFELVALMADSVLSDSPGP-----TWGRVTLVTFPAQTLLERG 111
      73 -----NMELQRMIA-AVDITSPREVFPRVAADMFSDGNEFMGRVALLFTYASKVLVKA 124
      112 -----PLVTARWKKGQFQPRLKEQEGDVARDQRLVALLSRHLMGQHRAMLQAQSGMDGF 166
      125 LCTKPELIRITMGM-----TLDFLRRLLG-----WIDQSGMGL 161
      167 CFFRTPPPLAFWRKQLVQAFISCLLTAFIYLMTRL 203
      162 LSYFGTP---TW-QTVTIFVAGVL-TASLTIWKKM 191

Db      162 LSYFGTP---TW-QTVTIFVAGVL-TASLTIWKKM 191

; RESULT 14
; US-08-470-865-13
; Sequence 13, Application US/08470865
; Patent No. 6586395
; GENERAL INFORMATION:
; APPLICANT: KIEFER, MICHAEL C.
; APPLICANT: BARR, PHILIP J.
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
; TITLE OR INVENTION: ENCODING THE PROTEINS AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,865
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/320,157
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.

```

```

: REGISTRATION NUMBER: 33,943
: REFERENCE/DOCKET NUMBER: 23647-20007.20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 813-5600
: TELEFAX: (415) 494-0792
: TELEX: 706141
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 192 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-470-865-13

Query Match 11.1%; Score 120; DB 4; Length 192;
Best Local Similarity 25.3%; Pred. No. 7.7e-06;
Matches 55; Conservative 29; Mismatches 71; Indels 62; Gaps 11

Db 10 TMAADPLRERTLLADYIGYCARBEG--TPBPA-PTSTPEAAVLRSAAARLRQIHRSFPSA 66
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
14 TSSEQIMKTGALLLGQFIQDPAGRWGGEAPELALDPVQDASTYKLSCLTKRIIGBELDS- 72
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
QY 67 YLGPGRNFELVALMADSVLSDSPG-----TWGRVVLVTVPAGTLRLRG 111
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 73 -----NMELQRMIA-AVDITDSPREVFRRVAADMFSQGNFNWGRVVALFYFASKLVYKA 124
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
QY 112 -----PLVTAAWKWKGFQPRLKEQBGVDARDCORIVALLISRLNGQHRAMVLAQGGWGF 166
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 125 LCTKVPFLIRITMGW-----TLDFLRRLRLIG-----WTIDQGGWGL 161
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
QY 167 CFFRTPEPLAFMRKQLVQAFLSCLLTAFIYIMTRL 203
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 162 LSYFGTP---TW--QTVITFVAGVL-TASLTIWKGM 191
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 15
US-09-155-327G-14
: Sequence 14, Application US/09155327G
: Patent No. 6790637
: GENERAL INFORMATION:
: APPLICANT: AMRAD Operations Pty Ltd
: TITLE OF INVENTION: A NOVEL MAMMALIAN GENE, bcl-2, BELONGS TO THE bcl-2
: TITLE OF INVENTION: FAMILY OF APOPTOSIS-CONTROLLING GENES
: FILE REFERENCE: 2096584
: CURRENT APPLICATION NUMBER: US/09/155,327G
: CURRENT FILING DATE: 1999-03-29
: PRIOR APPLICATION NUMBER: PN8965
: PRIOR FILING DATE: 1996-03-27
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 14
: LENGTH: 192
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-155-327G-14

Query Match 11.1%; Score 120; DB 4; Length 192;
Best Local Similarity 25.3%; Pred. No. 7.7e-06;
Matches 55; Conservative 29; Mismatches 71; Indels 62; Gaps 11

Db 10 TMAADPLRERTLLADYIGYCARBEG--TPBPA-PTSTPEAAVLRSAAARLRQIHRSFPSA 66
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
14 TSSEQIMKTGALLLGQFIQDPAGRWGGEAPELALDPVQDASTYKLSCLTKRIIGBELDS- 72
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
QY 67 YLGPGRNFELVALMADSVLSDSPG-----TWGRVVLVTVPAGTLRLRG 111
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 73 -----NMELQRMIA-AVDITDSPREVFRRVAADMFSQGNFNWGRVVALFYFASKLVYKA 124
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
QY 112 -----PLVTAAWKWKGFQPRLKEQBGVDARDCORIVALLISRLNGQHRAMVLAQGGWGF 166
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 125 LCTKVPFLIRITMGW-----TLDFLRRLRLIG-----WTIDQGGWGL 161
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
QY 167 CFFRTPEPLAFMRKQLVQAFLSCLLTAFIYIMTRL 203
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 162 LSYFGTP---TW--QTVITFVAGVL-TASLTIWKGM 191
: : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

```

Db :| |  
162 LSYFSTP---TW--QTVLIFVAGVL-TASLTIWKKM 191

Search completed: June 6, 2005, 10:45:12  
Job time : 43 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 6, 2005, 10:41:39 ; Search time 53 Seconds  
(without alignments)  
1330.542 Million cell updates/sec

Title: US-10-071-174A-2

Perfect score: 1084  
Sequence: 1 MVDQLRERTTADPLRERTE.....QAFSLCLTFAPIYMTRELL 204

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1084	100.0	204	14	US-10-071-174-2
2	701	66.7	130	9	US-09-912-599-7
3	328.5	30.3	141	9	US-09-764-847-800
4	328.5	30.3	141	14	US-10-092-154-800
5	153	14.1	30	14	US-10-071-174-19
6	128	11.8	24	14	US-10-071-174-13
7	121	11.2	191	14	US-10-196-793A-46
8	120	11.1	192	14	US-10-101-482-13
9	120	11.1	192	14	US-10-277-693A-9
10	120	11.1	192	14	US-10-306-878-5
11	120	11.1	192	15	US-10-421-285-10
12	120	11.1	331	9	US-09-033-525-2
13	120	11.1	331	17	US-10-640-668-2

14	118	10.9	23	9	US-09-912-599-14	Sequence 14, Appli
15	118	10.9	192	14	US-10-277-693A-2	Sequence 2, Appli
16	117	10.8	192	14	US-10-277-693A-8	Sequence 8, Appli
17	117	10.8	192	15	US-10-421-285-8	Sequence 8, Appli
18	117	10.8	192	16	US-10-451-467A-2	Sequence 7, Appli
19	117	10.8	193	9	US-09-925-674A-7	Sequence 6, Appli
20	117	10.8	193	16	US-10-450-366-6	Sequence 9, Appli
21	115	10.6	193	9	US-09-925-674A-9	Sequence 696, App
22	115	10.6	365	10	US-09-809-391-696	Sequence 696, App
23	115	10.6	365	10	US-09-882-171-696	Sequence 696, App
24	115	10.6	365	15	US-10-164-861-696	Sequence 696, App
25	114	10.5	21	14	US-10-071-174-6	Sequence 6, Appli
26	113	10.4	18	14	US-10-071-174-11	Sequence 31, Appli
27	112	10.3	237	16	US-10-659-705-5	Sequence 3, Appli
28	111	10.2	21	14	US-10-071-174-3	Sequence 3, Appli
29	111	10.2	192	14	US-10-277-693A-3	Sequence 3, Appli
30	111	10.2	239	15	US-10-148-953A-5	Sequence 5, Appli
31	110.5	10.2	152	14	US-10-158-769-2	Sequence 2, Appli
32	110	10.1	239	15	US-10-148-953A-3	Sequence 3, Appli
33	109	10.1	239	8	US-08-726-211-5	Sequence 5, Appli
34	109	10.1	239	10	US-09-993-420A-8	Sequence 8, Appli
35	109	10.1	239	14	US-10-277-693A-10	Sequence 10, Appli
36	109	10.1	239	14	US-10-141-618-12	Sequence 12, Appli
37	109	10.1	239	14	US-10-053-645A-21	Sequence 21, Appli
38	109	10.1	239	15	US-10-387-961A-5	Sequence 5, Appli
39	109	10.1	239	15	US-10-003-632C-1	Sequence 1, Appli
40	109	10.1	239	15	US-10-003-632C-3	Sequence 3, Appli
41	109	10.1	239	15	US-10-003-632C-10	Sequence 10, Appli
42	109	10.1	239	15	US-10-003-632C-13	Sequence 13, Appli
43	109	10.1	239	15	US-10-148-953A-1	Sequence 1, Appli
44	109	10.1	239	15	US-10-148-953A-2	Sequence 2, Appli
45	109	10.1	239	15	US-10-148-953A-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1									
US-10-071-174-2									
; Sequence 2, Application US/10071174									
; Publication No. US20030176671A1									
; GENERAL INFORMATION:									
; APPLICANT: REED, JOHN C.									
; APPLICANT: KE, NING									
; APPLICANT: GODDIX, ADAM									
; TITLE OF INVENTION: APOPTOSIS MODULATOR BCL-2 AND METHODS FOR MAKING AND									
; FILE REFERENCE: 087102-0272558									
; CURRENT APPLICATION NUMBER: US/10/071,174									
; CURRENT FILING DATE: 2002-02-07									
; PRIOR APPLICATION NUMBER: 60/267,166									
; PRIOR FILING DATE: 2001-02-07									
; NUMBER OF SEQ ID NOS: 36									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 2									
; LENGTH: 204									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-071-174-2									
Query Match									
Best Local Similarity 100.0%; Score 1084; DB 14; Length 204;									
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MVDQLRERTTADPLRERTELLADYLCAREPCTPPAPSTPAAVLRSAARLRQIH	60						
DB	1	MVDQLRERTTADPLRERTELLADYLCAREPCTPPAPSTPAAVLRSAARLRQIH	60						
QY	61	RSFSAAYIGYGNRRELVAMADSVLSDSPPTWRRVTVTLTPAGTLLERGLTVARKK	120						
DB	61	RSFSAAYIGYGNRRELVAMADSVLSDSPPTWRRVTVTLTPAGTLLERGLTVARKK	120						
QY	121	WQFGRLEKRGEDVARDQRTVALLSRLMGQHRAMLQAGGMDGFCFFRTPTPLAFWR	180						

```
Db 121 WFFQRLKEQEGDVARDCCRLVALLSRLMGQHRAMLQAGGMDGFCHEFRTPFLAFWR 180
Qy 181 KOLVQAFSLCLTTAFIYLMTRL 204
Db 181 KOLVQAFSLCLTTAFIYLMTRL 204

RESULT 2
US-09-912-599-7
; Sequence 7, Application US/09912599
; Patent No. US20020106731A1
; GENERAL INFORMATION:
; APPLICANT: Ruden et al.
; TITLE OF INVENTION: Bcl-2-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO44P1
; CURRENT APPLICATION NUMBER: US/09/912,599
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: PCT/US01/03080
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/179,487
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/180,697
; PRIOR FILING DATE: 2000-02-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-912-599-7

Query Match 64.7%; Score 701; DB 9; Length 130;
Best Local Similarity 100.0%; Pred. No. 7.5e-64;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 FELVALMADSVLSDSPGPTMGVTVTLVTFAGTLLERGPVLTARMKMGFQPRLKQEGDV 134
Db 1 FELVALMADSVLSDSPGPTMGVTVTLVTFAGTLLERGPVLTARMKMGFQPRLKQEGDV 60

Qy 135 ARDCORVALLSRLMGQHRAMLQAGGMDGFCHEFRTPFLAFWRKOLVQAFSLCLTT 194
Db 61 ARDCORVALLSRLMGQHRAMLQAGGMDGFCHEFRTPFLAFWRKOLVQAFSLCLTT 120

Qy 195 AFYLMTRL 204
Db 121 AFYLMTRL 130

RESULT 3
US-09-764-847-800
; Sequence 800, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAMM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 800
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

```
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (86)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (108)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-800

Query Match 30.3%; Score 328.5; DB 9; Length 141;
Best Local Similarity 67.0%; Pred. No. 1.5e-25;
Matches 67; Conservative 1; Mismatches 11; Indels 21; Gaps 2;

Qy 90 PGPTMGVTVTLVTFAGTLLERGPVLTARMKMGFQPRLKQEGDVARDCCORVALLSRL 149
Db 40 PGPTMGVTVTLVTFAGTLLERGPVLTARMKMGFQPRLKQEGDVARDCCORVALLSRL 99

Qy 150 MGOHRA-----WTQAGGMDG-FCH 168
Db 100 MGOHRRPGCKLRAGMAFVTSSGPFFHMLFGENSMRGLFCH 139

RESULT 4
US-10-092-154-800
; Sequence 800, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 800
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (86)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (108)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-092-154-800

Query Match 30.3%; Score 328.5; DB 14; Length 141;
Best Local Similarity 67.0%; Pred. No. 1.5e-25;
```

Matches 67; Conservative 1; Mismatches 11; Indels 21; Gaps 2;

QY 90 PGPWGRVTVLTFTAGTLLERGLVTARWKKGQPRLEKGGDVARDQRLVALLSRL 149  
Db 40 PRPHGRVTVLTFTAGTLLERGLVTARWKKGQPRLEKGGDVARDQRLVALLSRL 99  
QY 150 MGOHRA-----MTAOQGMDG-FCH 168  
Db 100 MGOHRA-----MTAOQGMDG-FCH 139

RESULT 5  
US-10-071-174-19  
; Sequence 19, Application US/10071174  
; Publication No. US20030176671A1  
; GENERAL INFORMATION:  
; APPLICANT: REED, JOHN C.  
; APPLICANT: KE, NING  
; APPLICANT: GODZIK, ADAM  
; TITLE OF INVENTION: APOPTOSIS MODULATOR BCL-B AND METHODS FOR MAKING AND  
; FILE REFERENCE: 087102-0272558  
; CURRENT FILING DATE: 2002-02-07  
; PRIOR APPLICATION NUMBER: 60/267,166  
; PRIOR FILING DATE: 2001-02-07  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-071-174-19

Query Match 14.1%; Score 153; DB 14; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 MADSVLSDSPGPWGRVTVLTFTAGTLLER 110  
Db 1 MADSVLSDSPGPWGRVTVLTFTAGTLLER 30

RESULT 6  
US-10-071-174-13  
; Sequence 13, Application US/10071174  
; Publication No. US20030176671A1  
; GENERAL INFORMATION:  
; APPLICANT: REED, JOHN C.  
; APPLICANT: KE, NING  
; APPLICANT: GODZIK, ADAM  
; TITLE OF INVENTION: APOPTOSIS MODULATOR BCL-B AND METHODS FOR MAKING AND  
; FILE REFERENCE: 087102-0272558  
; CURRENT FILING DATE: 2002-02-07  
; PRIOR APPLICATION NUMBER: 60/267,166  
; PRIOR FILING DATE: 2001-02-07  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-071-174-13

Query Match 11.8%; Score 128; DB 14; Length 24;  
Best Local Similarity 100.0%; Pred. No. 6.6e-06;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LRRETELLADYLGVCAREPGTPE 38  
Db 1 LRRETELLADYLGVCAREPGTPE 24

RESULT 7  
US-10-196-793A-46  
; Sequence 46, Application US/10196793A  
; Publication No. US20030086919A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSENBLUM, MARTIN G.  
; APPLICANT: LIU, YIYING  
; TITLE OF INVENTION: THERAPEUTIC AGENTS COMPRISING PRO-APOPTOTIC PROTEINS  
; FILE REFERENCE: CLFR:01205  
; CURRENT FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: 60/360,361  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/332,886  
; PRIOR FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: 60/306,091  
; PRIOR FILING DATE: 2001-07-17  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-196-793A-46

Query Match 11.2%; Score 121; DB 14; Length 191;  
Best Local Similarity 27.4%; Pred. No. 0.00045;  
Matches 58; Conservative 29; Mismatches 73; Indels 52; Gaps 12;

QY 10 TMADPLRRTRELLADYLGVCAREPG--TPRPA-PTTEAAVLRRAARLQIHSPSA 66  
Db 13 TSSQIMTGTALLQGFQDRAGRWGSEAPALADPVQDASTKLSCLKRIGDELDS- 71  
QY 67 YLGYGNRFELVALMADSVLSDSPG-----TWGRVTVLTFTAGTLLER 111  
Db 72 -----NMELORMIA-AVDYDSPREVFRAADMFSDGNRWGKVALLFFASLVYKA 123  
QY 112 PLVTARWKKGQPRLEKGGDVARDQRLVALLSSRLMGHRAWLQAOQGMDFCHFR 171  
Db 124 -LVTK-----VPEL-----IRTMGTLDPLREIRLG-----WIDQGGMDGLSYFG 165  
QY 172 TPPLAFWRKQLVQAFSLCLTTATAYLWTRL 203  
Db 166 TP---TW---QVTITFVAGVL-TASLTIWKKM 190

RESULT 8  
US-10-101-482-13  
; Sequence 13, Application US/10101482  
; Publication No. US2003008837A1  
; GENERAL INFORMATION:  
; APPLICANT: KIEFER, MICHAEL C.  
; APPLICANT: KIEFER, PHILIP J.  
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/101,482  
; FILING DATE: 18-Mar-2002

```

? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/320,157
? FILING DATE: 07-OCT-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: LEHNHARDT, SUSAN K.
? REGISTRATION NUMBER: 33,943
? REFERENCE/DOCKET NUMBER: 23647-20007.20
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 813-5600
? TELEFAX: (415) 494-0792
? TELEX: 706141
? INFORMATION FOR SEQ ID NO: 13:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 192 amino acids
? TYPE: amino acid
? STRADEDNESS: single
? TOPOLOGY: linear
? SEQUENCE DESCRIPTION: SEQ ID NO: 13:
? US-10-101-482-13

```

SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-10-101-482-13

Query Match	11.1%	Score 120;	DB 14;	Length 192;
	11.1%	Score 120;	DB 14;	Length 192;

Best Local Similarity 25.3%; Pred. No. 0.00057;  
Matches 55; Conservative 29; Mismatches 71; Indels 62; Gaps 11,

Oy 10 TMADPLRRETELLIADYLGCAREPG-TPEPA-PSTPEAAVLNSAARLRQIHHSFFSA 66  
 : : : : :  
 Db 14 TSSOIMKTCGALLGGFIODRAGRWGEAPELADVPVPOASTKTLSECCIRIGDELDLS-72  
 : : : : :  
 : : : : :

```
QY      67 YLGIVGNRFELVALMADSVLSDSPG-----TWGRVVTLVTPAGTLBERG   111
          ||::|:||||||| | | | | | | | | | | | | | | | | | | | | |
Db      73 -----INMEIQRMIA-AVDITDPPREVFPRVAADMFDGDNFMGWGVALLFYFASLTVLKA   124
```

```

QY      112  ----PLVTAAWKKVGFQRLKEQGDVARDCCQLVALISLSRLMGQRAMILQAQSGMDGF 166
          | : | | | | |
Db      125  LCTVPELIRITMGV-----TLDFLRERLLG---WIQDQSGMDGL 161

```

```
Qy 167 CHFFRTFFPLAEWRKQIVQAFSLCUIITTAIFYLWTRL 203
      :|||_|_|:::|_|:::
Db 162 LSYFGTF---TW--QVTIFVAGVL-TASLTIVKKKG 191
```

RESULT 9  
US-10-277-693A-9  
Sequence 9 Application US/10277693A

Publication No. US20030096367A1  
GENERAL INFORMATION:  
APPLICANT: Kormsmeier, Stanley J.  
TITLE OF INVENTION: Cell Death Agonists

FILE REFERENCE: 56029/36280  
CURRENT APPLICATION NUMBER: US/10/277,693A  
CURRENT FILING DATE: 2002-10-22  
PRIOR APPLICATION NUMBER: 09/379,820

; PRIOR FILING DATE: 1999-08-24  
 ; PRIOR APPLICATION NUMBER: 08/112,208  
 ; PRIOR FILING DATE: 1993-08-26  
 ; PRIOR APPLICATION NUMBER: 08/856,034  
 ; PRIOR FILING DATE: 1997-05-14

```

; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1.1
; SEQ ID NO 9
; LENGTH: 192

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-277-693A-9

```

Query Match	11.1%;	Score 120;	DB 14;	Length 192;
Best Local Similarity	25.3%;	Pred. No. 0.00057;		
Matches 55;	Conservative 29;	Mismatches 71;	Indels 62;	Gaps 11

**Dy** 10 TMADPLRRTELLADLYGTCAREG--TPPEA-PTSEPAVLKRSAAARLRQIHRSFSA 66  
::: ||| :: | | : : :  
**Db** 14 TTSSEQIMKTGALLIGFTIQDRAGRMGEAPLALPVPVDPASTKKLSECLKRIGBELDS- 72

```

QY 67 YLGYGNRFELVALMADSVLSDSPB-----TWGRVTLVTFAGTLLBEG 111
DB 73 -----NMELQRMIA-AVDITSPREVEFFRVAADMFSDGNFMGRVVALFTFYSKVLVKA 124
QY 112 -----PLVTATMKKKGFQPRLKEGSDVARDCQRLVALLSRRLMGORAMLAOGSDMF 166
DB 125 LCTKVPBELIRITMGN-----TLDFLRERKLG--WIDOGSGMDGL 161
QY 167 CHFFRTPEPLAFWRKOLVOAFPLSCLLTFAFIYLRML 203
DB 162 LSYFCTP---TW-QIVTIFVAGVL-TASLTITMKM 191

```

RESULT 10  
US-10-306-878-5  
Sequence 5 Application US/10306878

```

/ Signature of Applicant: 067455000
/ Publication No. US20030175819A1
/ GENERAL INFORMATION:
/ APPLICANT: Reed, John C.
/ APPLICANT: Guo, Bin

```

```

; TITLE OF INVENTION: Methods for Identifying Modulators of
;
; TITLE OF INVENTION: Apoptosis
;
; FILE REFERENCE: P-LJ 5535
;
; CURRENT APPLICATION NUMBER: US/10/306,878
;
; CURRENT FILING DATE: 2002.11.27

```

```

; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: US 60/334,149
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 5
;
; LENGTH: 192
;
; TYPE: prt
;
; ORGANISM: Homo sapiens
;

```

US-10-306-878-5

	531	Colset Value	231	Min Max	72	Index	021	Copy	2
QY	10	TMADPLRTELLADLYG	CAREBG-	-TPBA	-PSTPEAAVL	RSAAARL	ROIHS	SFPSA	66
		:::	:::	:::	:::	:::	:::	:::	
Db	14	TSSEIMKTGALLGFI	QDRAGRMGEAP	ELALDP	VPDAS	TKUSIECL	KRGDELD-		72

```

67 YLGVGNRFELVALMADSVLSDSPG-----TWGRVITLVTPAGTILRRG 111
    ||::||::||| ||||| ||:::
QY -----IMELQRMIA-AVDITDSPREVEFRRVADMFSDGNFMGRGVALLFYFSKVLVKA 124
Db 73 -----

```

```

QY      112 -----PLVTARWKXKGFQFRLKEGSDVARDQCRVALLSSRLMGHRAWLDAQGSMDFG 166
          | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D5      125 LCTKVPBLIRITMGV-----TLDPLRERLDG---WIDQGGSMDFG 161

```

```

QY      167 CHEFRTPPLAFWRKQLVQAFISCLLTATATYILWTRL 203
      : || : : : : : :
Db      162 LSIFGTP---TW--QIVTIFAGVL-TASLTIMKKM 191

```

**RESULT 11**  
US-10-421-285-10  
; Sequence 10, Application US/10421285  
Publ:act:ca No: us200406282ca1

! PUBLIC INFORMATION NO. US20040035358A1  
!  
! GENERAL INFORMATION:  
!  
! APPLICANT: MayerKuckuk, Phillip  
!  
! APPLICANT: Banerjee, Debabrata  
!  
! APPLICANT: Bertino, Joseph R.

```

; TITLE OF INVENTION: Method for Modulating the Production of a Selected Protein In
;
; TITLE OF INVENTION: V1V0
; FILE REFERENCE: MSK-P-053
; CURRENT APPLICATION NUMBER: US/10/421,285

```

```

; CURRENT FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US 60/375,250
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 15

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 192
; TYPE: PRT
; ORGANISM: human
US-10-421-285-10

```

Query Match	11.1%	Score 120;	DB 15;	Length 192;
Best Local Similarity	25.3%	Pred. No. 0.00057;		
Matches	55;	Conservative	29;	Mismatches 71;
			Indels	62;
			Gaps	11.

[illegible]

```

RESULT 12
US-09-033-525-2
; Sequence 2, Application US/09033525
; Patent No. US20020090374A1
; GENERAL INFORMATION:
; APPLICANT: Yarkoni, Shai
; APPLICANT: Ben-Yehudah, Ami
; APPLICANT: Azari, Yehudit
; APPLICANT: Agellan, Rami
; APPLICANT: Belostotsky, Ruth
; APPLICANT: Lorberbaum-Galaki, Haya
; TITLE OF INVENTION: CHEMIC PROTEINS WITH CELL-TARGETING
; TITLE OF INVENTION: SPECIFICITY AND APOPTOSIS-INDUCING ACTIVITIES
; FILE REFERENCE: 9457-009-999
; CURRENT APPLICATION NUMBER: US/09/033,525
; CURRENT FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
; OS-09-033-525-2

```

[illegible]

```

RESULT 13
US-10-640-668-2
; Sequence 2, Application US/10640668
; Publication No. US20050079154A1
; GENERAL INFORMATION:
; APPLICANT: Yarkoni, Shai
; APPLICANT: Ben-Yehudiah, Ahmi
; APPLICANT: Azar, Yehudich
; APPLICANT: Ageljan, Rami
; APPLICANT: Belostotsky, Ruth
; APPLICANT: Lorberboun-Galski, Haya
; TITLE OF INVENTION: CHIMERIC PROTEINS WITH CELL-TARGETING
; TITLE OR INVENTION: SPECIFICITY AND APOPTOSIS-INDUCING ACTIVITIES
; FILE REFERENCE: 9457-009-999
; CURRENT APPLICATION NUMBER: US/10/640,668
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: US/09/033,525
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-640-668-2

```

[illegible]

```

RESULT 14
US-09-912-599-14
; Sequence 14, Application US/09912599
; Patent No. US20020106731A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Bcl-2-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO44P1
; CURRENT APPLICATION NUMBER: US/09/912,599
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: PCT/US01/03080
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/179,487
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/180,697
; PRIOR FILING DATE: 2000-02-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-912-599-14

Query Match      10.9%; Score 118; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 6; Ee-05;

```

Matches	23; Conservative	0; Mismatches	0; Indels	0; Gaps
Qy	181 KQLVQAFISCLTTAFIYLWTRL	203		
Db	1 KQLVQAFISCLTTAFIYLWTRL	23		

```

RESULT 15
US-10-277-693A-2
: Sequence 2, Application US/10277693A
: Publication No. US20030096367A1
: GENERAL INFORMATION:
: APPLICANT : Kosemeyer, Stanley J.
: TITLE OF INVENTION: Cell Death Agonists
: FILE REFERENCE: 56029/36280
: CURRENT APPLICATION NUMBER: US/10/277,693A
: CURRENT FILING DATE: 2002-10-22
: PRIOR APPLICATION NUMBER: 09/379, 820
: PRIOR FILING DATE: 1999-08-24
: PRIOR APPLICATION NUMBER: 08/112, 208
: PRIOR FILING DATE: 1993-08-26
: PRIOR APPLICATION NUMBER: 08/856, 034
: PRIOR FILING DATE: 1997-05-14
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: PatentIn Version 3.1
: SEQ ID NO 2
: LENGTH: 192
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-277-693A-2

```

[illegible]

Search completed: June 6, 2005, 10:46:10  
Job time : 54 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: June 6, 2005, 10:37:59 ; Search time 39 Seconds

(without alignments)  
503.288 Million cell updates/sec

Title: US-10-071-174A-2

Sequence: 1 MVDLRRTMTADPLRERTE.....QAFSLCLTAPFYIMTRL 204

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	193.5	17.9	177	2	SS4778
2	123	11.3	192	2	D47538
3	120	11.1	192	2	A47538
4	113	10.4	143	2	I38921
5	109	10.1	239	1	TVHUA1
6	108	10.0	350	2	A47476
7	105.5	9.7	133	2	I53295
8	101.5	9.4	172	2	I49449
9	101.5	9.4	236	2	JC7383
10	101.5	9.4	236	2	I67432
11	100.5	9.3	233	2	A37332
12	99.5	9.2	233	2	B47537
13	99.5	9.2	233	2	I49056
14	99.5	9.2	233	2	IS1761
15	99.5	9.2	236	2	I53744
16	98.5	9.1	214	2	I49057
17	98	9.0	227	2	JR0203
18	95.5	8.8	205	1	TVHUB1
19	95.5	8.8	255	2	UC7567
20	95	8.8	236	1	TVMSA1
21	94.5	8.7	233	2	I67431
22	94	8.7	232	2	S24390
23	91	8.4	199	1	TWMSB1
24	85.5	7.9	380	2	A75500
25	84.5	7.8	192	2	A10268
26	84	7.7	216	2	B37332
27	83	7.7	272	2	T46359
28	82.5	7.6	658	2	T46359
29	82	7.6	190	2	A47537

30	82	7.6	334	2	T17213	hypothetical prote
31	81.5	7.5	175	2	I39055	Bcl-2 related - hu
32	81	7.5	915	2	T09575	smoothelin - human
33	80.5	7.4	179	2	UC7255	Bax-delta protein
34	80	7.4	209	2	B82241	preprotein translo
35	79.5	7.3	522	1	S01927	nif-specific regul
36	78.5	7.2	200	1	NNBB2D	anthranilate synth
37	78	7.2	218	2	B47538	bcl-2-associated p
38	78	7.2	515	2	H75579	hypothetical prote
39	78	7.2	535	2	T37508	hypothetical prote
40	78	7.2	1632	2	C70752	probable ctcp1 prot
41	77	7.1	365	2	T34759	oligopeptide ABC t
42	77	7.1	462	2	H70593	hypothetical prote
43	77	7.1	543	2	T34681	probable secreted
44	76.5	7.1	354	2	B87635	efflux system prot
45	76.5	7.1	386	2	T36169	hypothetical prote

## ALIGNMENTS

```

RESULT 1
S54778
NR-13 protein - qual
C:Species: Coturnix coturnix (qual)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S54778
R:Gillet, G.; Guerin, M.; Trembleau, A.; Brun, G.
EMBO J. 14, 1372-1381, 1995
A>Title: A BCL-2-related gene is activated in avian cells transformed by the Rous sarcoma
A:Reference number: S54778; MUID:95246730; PMID:7729415
A:Accession: S54778
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-177 <GIL>
A:Cross-references: EMBL:X84418; NID:g790983; PIDN:CAAS9136.1; PID:g790984

Query Match          17 9%; Score 193.5; DB 2; Length 177;
Best Local Similarity 33.1%; Pred.No.2.4e-11;
Matches 57; Conservative 16; Mismatches 56; Indels 43; Gaps 6;

QY 11 MADPLRERTELLADYLCYCARPEPTPEAPSTPEAAVLRSAARLRQIHRSFPAVY-- 68
DB 1 MPGLIKERTALLLEVPQH--RAGGAALPPSAT--AAELRRAAAALBRERPFRRSCAPL 56
QY 69 --GYPNRFELVALMADSVLSDSPGPTWGRVTVLTFRGTL-----LEKGLVTARW 118
DB 57 ARAEPREAAALLRKVAQLRTDG-GLNWGRLLALVVFAGTLAAALAESACEBGP----- 109
QY 119 KKNQFQPLKQEGDVANDCQRLVALLSRLMGQRARLQAGCGWDCRCHFP 170
DB 110 -----SRLAAALTYVLAERQGEWMBEHBGWDGCRFP 141

RESULT 2
D47538
bcl-2-associated protein x - mouse
N:Alternate names: BAX; programmed cell death membrane protein x
C:Species: Mus musculus (house mouse)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 28-Jul-2003
C:Accession: D47538
R:Olvera, Z.N.; Millman, C.L.; Korsmeyer, S.J.
Cell 74, 609-619, 1993
A>Title: Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates p
A:Reference number: A47538; MUID:93364978; PMID:8358790
A:Accession: D47538
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-192 <OLV>
A:Cross-references: GB:I22472
C:Genetics:
A:Gene: bax
C:Superfamily: bcl apoptosis regulator, inhibitory type

```

	Query Match	11.3%:	Score 123;	DB 2;	Length 192;
	Best Local Similarity	26.7%:	Pred. No. 0.00016;		
	Matches	54;	Conservative	23;	Mismatches 63; Indels 62; Gaps 11;
QY	23	LADYLGYCAREPGTPE--PAPSTPEAAVLRSAARLROIHRFSFYSALGYDGNRDELVALM	81		
	:	:         :	:	:	:
Db	31	IDRAGRNAGE--TPELTLBEPQDASTKSLSECCRRIGDELDG-----NNELGRMI	80		
QY	82	ADSVLSGSGP-----TWGRVYLVTFAGTLLBEG-----PLYFAARKKW	121		
	:	:	:	:	:
Db	81	AD-VDTDSPREVFPRVAADMFDGNGFNMGVRVALFYFPASKVLVALCTKYPELTRITMGW	139		
QY	122	GFOPLKEQBEDVARDCORIVALLISRLMGCHRAMLAQOGSMDFCFHFRFFFLAEMRK	181		
	:	:         :	:	:	:
Db	140	-----TLDFLRERLTG----WIDQGGMGBGLTSFGTR-----TW--	170		
QY	182	QLVOAFSLSCLLTFAFIYLMTRL	203		
	:	:	:	:	:
Db	171	QTVTIFVAGVL-TASLTIMKKM	191		

```

RESULT 3
A47538
bcl-2-associated protein x, alpha splice form - human
N/Alternate names: BAX; programmed cell death membrane protein x alpha
C/Species: Homo sapiens (man)
C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C/Accession: A47538
R/Oltvai, Z.N.; Millman, C.L.; Korsmeyer, S.J.
Cell 74, 609-619, 1993
A/Title: Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates F
A/Reference numbers: A47538; MUID:93364978; PMID:8358790
A/Accession: A47538
A/Molecule type: mRNA
A/Residues: 1-192 <OLT>
A/Cross-references: UNIPROT:Q07812; GB:L22473; NID:G388165; PIDN:AAA03619.1; PID:G388166
A/Note: the amino end of the mature protein is blocked
C/Genetics:
A/Gene: GDB:BAX
A/Cross-references: GDB:228082; OMTM:600040
A/Map position: 19q13.3-19q13.4
C/Superfamily: bcl apoptosis regulator, inhibitory type
C/Keywords: alternative splicing; blocked amino end; heterodimer; homodimer; transmembrane
#172-191/pomam: transmembrane #status predicted <TMM>

```

[illegible]

```

RESULT 4
138921
bcl-2-associated protein x, delta splice form - human
N:Alternate names: BAX; programmed cell death membrane protein x delta
C:Species: Homo sapiens (man)
C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 09-Jul-2004

```

C:Accession: I38921  
R:Appt. S.S.; Matrei. M.G.; Olsen, B.R.  
Genomics 26, 592-594, 1995  
A:Title: Mapping of the human BAX gene to chromosome 19q13.3-q13.4 and isolation of a cDNA  
A:Reference number: I38921, MUID:95331797, PMID:7607665  
A:Accession: I38921  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-143 <RES>  
A:Cross-references: UNIROT:P55269; EMBL:U19599; NID:g841237; PIDN:AACS0142.1; PID:g841237  
C:Genetics:  
A:Gene: GDB:BAX  
A:Cross-references: GDB:228082; OMIM:600040  
A:Map position: 19q13.3-19q13.4  
C:Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 10.4%; Score 113; DB 2; Length 143;  
Best Local Similarity 28.6%; Pred. NO. 0.0012;  
Matches 38; Conservative 42; Indels 36; Gaps 7;

[illegible]

RESULT 5

transforming protein bcl-2, splice form alpha - human

TVHVAL

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1988 #sequence revision 07-Jun-1996 #text change 09-Jul-2004

C:Accession: C37332; A29409; S02452; A24428; A27622; B27622

R:Enguchi, Y.; Ewert, D.L.; Tsujimoto, Y.

N: Nucleic Acids Res. 20, 4187-4192, 1992

A:Title: Isolation and characterization of the chicken bcl-2 gene: expression in a variety of tissues

A:Accession: C37332

A:Reference number: A37332; MUID:92375724; PMID:1508772

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Stem: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-239 <EGU>

A:Cross-references: UNIPROT:P10415

A>Note: this report is a correction

R:Tsujimoto, Y.; Croce, C.M.

Proc. Natl. Acad. Sci. U.S.A. 83, 5214-5218, 1986

A:Title: Analysis of the structure, transcript, and protein products of bcl-2, the gene for the B-cell lymphoma

A:Reference number: A29409; MUID:86235760; PMID:3523487

A:Accession: A29409

A:Molecule type: mRNA

A:Residues: 1-95, 'A', '97-109, 'G', '111-236, 'S', '238-239 <TSU>

A:Cross-references: GB:M13994; NID:G179366; PIDN:AAA51813.1; PID:G179367

A>Note: this sequence has been corrected in reference A37332

R:Seco, M.; Jaeger, U.; Hockett, R.D.; Gruninger, W.; Bennett, S.; Goldman, P.; Korsmeyer, S.J.

EMBO J. 7, 123-131, 1988

A:Title: Alternative promoters and exons, somatic mutation and deregulation of the Bcl-2 gene

A:Reference number: S02452; MUID:88196071; PMID:2834197

A:Accession: S02452

A:Molecule type: mRNA

A:Residues: 1-239 <SRT>

R:Cleary, M.L.; Smith, S.D.; Sklar, J.

Cell 47, 19-28, 1986

A:Title: Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin heavy chain enhancer

A:Reference number: A24428; MUID:87002488; PMID:2875799

A:Accession: A24428

A:Molecule type: mRNA

A:Residues: 1-58, 'T', '60-116, 'R', '118-239 <CLB>

A:Cross-references: GB:M14745; NID:G179370; PIDN:AAA35591.1; PID:G179371



R.Hua, C.; Zorn, S.; Jensen, J.P.; Coupland, R.W.; Ko, H.S.; Wright, J.J.; Bakshi, A.  
Oncogene Res. 2, 263-275, 1988  
A:Title: Consequences of the t(14;18) chromosomal translocation in follicular lymphoma:  
A:Reference number: A27622; MUID:88217344; PMID:3285301  
A:Accession: A27622  
A:Molecule type: mRNA  
A:Residues: 1-58, 'T', 60-239 <HUA>  
A:Accession: B27622  
A:Molecule type: DNA  
A:Residues: 1-6, 'S', 8-58, 'T', 60-128, 'C', 130-239 <HUA2>  
A>Note: the sequence was determined from the germ-line gene  
C:Comment: Constitutive expression of BCL2 following t(14;18) chromosomal translocation  
C:Genetics:  
A:Gene: GDB:BCL2  
A:Cross-references: GDB:119031; OMIM:151430  
A:Map position: 18q21.3-18q21.3  
C:Function:  
A:Description: blocks apoptosis in hematopoietic cells  
C:Superfamily: bcl apoptosis regulator, inhibitory type  
C:Keywords: alternative splicing; apoptosis; B-cell lymphoma; follicular lymphoma; proto

Query Match 10.1%; Score 109; DB 1; Length 239;  
Best Local Similarity 23.5%; Pred. No. 0.0052;  
Matches 52; Conservative 20; Mismatches 65; Indels 84; Gaps 11;

28 GYCARPG-TPEPA-----PSTPEA-----VLRSA----- 52  
Db 47 GIPSSQPGHTHPAASRDPAVTSPLQTPAAGAAAPALSVPPVHLLTLOQADDSR 106  
Qy 53 -----AARLRQIHRSFSAVLYGPGNRFELVALMADSVLSDSPGTVGVVTLVTFAGT 106  
Db 107 RYRDPFAMSSQHLTPTFA-----RGRF---ATVEBELFRD--GVNMGRIVAFFFGV 156  
Qy 107 LL-----ERGPLVTARMKKMGFQRLKEQSDVARDCORVALVLSRLMGH-RAMTQA 159  
Db 157 MCVSVNNEPSLYDN-----IALMWTETLNHHTWIQD 191  
Qy 160 QGMDGCFHFRTPF-PLAFMRKQLVQAFSLGLTLTAPITL 199  
Db 192 NGMDAFVELYGPSMRPLPFSWSLKTLLSLALVACITL 232

RESULT 6  
A47476  
BCL2 homolog MCL1 - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C:Accession: A47476  
R:Kozopas, K.M.; Yang, T.; Buchan, H.L.; Zhou, P.; Craig, R.W.  
Proc. Natl. Acad. Sci. U.S.A. 90, 3516-3520, 1993  
A:Title: MCL1, a gene expressed in programmed myeloid cell differentiation, has sequence  
A:Reference number: A47476; MUID:93234528; PMID:7682708  
A:Accession: A47476  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-350 <KOZ>  
A:Experimental source: ML-1 myeloid cell leukemia  
A>Note: sequence extracted from NCBI Backbone (NCBI:129734)  
C:Keywords: transmembrane protein

Query Match 10.0%; Score 108; DB 2; Length 350;  
Best Local Similarity 23.8%; Pred. No. 0.01;  
Matches 49; Conservative 20; Mismatches 71; Indels 66; Gaps 8;

Qy 13 DPLRETELL-LADYLGACAREPGTPEPAATP----- 44  
Db 132 EPLGRPAVLPLLELVGSSGNNSTSDGLSPFPAAEEBDELVQSLIISRYLREQAT 191  
Qy 45 -----EAYVRSAAARLQI-----HNSFSAVLYGPG--NRFELVALMADSV 85  
Db 192 GAKDTKPMGRSGATSRKALETLRRVGDVORHETVFGMLRKLDIKNEDVKSLSRYMI 251  
Qy 86 LSDSPGPT-WGRVVTLVTFAGTLERGPLVTARMKKMGFQRLKEQSDVARDCORVAL 144

Db 252 HVFSDGVTNMGIVTLISF-----GAFVAKHLK-----TINQESCI-----EPLAES 293  
Qy 145 LSSRLMGCHRAMLQAGQMDGFCHEF 170  
Db 294 ITDVLVRRTRDMLVQRQMDGFVEPF 319

RESULT 7  
153295  
bcl-2 associated protein x - rat (fragment)  
N:Alternate names: BAX; programmed cell death membrane protein x  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 09-Jul-2004  
C:Accession: 153295  
R:Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.  
Endocrinology 136, 233-241, 1995  
A:Title: Expression of members of the bcl-2 gene family in the immature rat ovary: equin  
onstitutive bcl-2 and bcl-xlong messenger ribonucleic acid levels.  
A:Reference number: 153295; MUID:95129487; PMID:7828536  
A:Accession: 153295  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-133 <RES>  
A:Cross-references: UNIPROT:Q63690; EMBL:U32098; NID:g975669; PIDN:AAA75200.1; PID:g9758

Query Match 9.7%; Score 105.5; DB 2; Length 133;  
Best Local Similarity 27.2%; Pred. No. 0.0058;  
Matches 44; Conservative 16; Mismatches 43; Indels 59; Gaps 9;

Qy 36 TPE-PAPSTPEAAVIRSAARLRQIHRSFSAVLYGPGNRFELVALMADSVLSDSPG-- 92  
Db 6 TPELTLEQPPQDASTKSLSECLRRIGDEL-----DNNMELQRMID-VDTSPREVF 56  
Qy 93 -----TWGRVYTLVTFAGTLERGPLVTARMKKMGFQRLKEQSDVARDCQ 139  
Db 57 FRVAADMFADGNFMGRVVALFYFASKLV-----IKAL-----CT 91

Qy 140 RLVALSSRLMG-----QHR--AMLQAGQMDGFCHEFRTP 173  
Db 92 KVPBLIRT-IMGWTLDFLRRLFWIQDQSGMDGLSTFGTP 132

RESULT 8  
149449  
hemopoietic-specific early-response protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: 149449  
R:Linn, E.Y.; Orloffsky, A.; Berger, M.S.; Prytcowsky, M.B.  
J. Immunol. 151, 1979-1988, 1993  
A:Title: Characterization of A1, a novel hemopoietic-specific early-response gene with s  
A:Reference number: 149449; MUID:93346743; PMID:8345131  
A:Accession: 149449  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-172 <RES>  
A:Cross-references: UNIPROT:Q07440; GB:U16462; NID:g293273; PIDN:AAA16886.1; PID:g293274

Query Match 9.4%; Score 101.5; DB 2; Length 172;  
Best Local Similarity 23.6%; Pred. No. 0.019;  
Matches 39; Conservative 31; Mismatches 66; Indels 29; Gaps 8;

Qy 22 LLADYLGACAREPGTPEPAATPAAVIRSAARLRQIHRSFSAVLYGPGNRFELVALM 81  
Db 12 LAEHYLOVITLOVPAP-ESAPQACRVLOVAFSVQKEVKLKSLYDDPHVESITDKRTI 70  
Qy 82 ADVSL--DSPGTVGRVVTLVTFAGTLERGPLVTARMKKMGFQRLKEQSDVARDC 137  
Db 71 FNQVMEKRFEDGIINMGRIVITIFAFGVTLKKLP-----QEQ---IALDV 112



**C:Keywords:** alternative splicing; apoptosis  
**F1-1-233/Product:** apoptosis regulator bcl-xl #status predicted <MAT>  
**F1-125,189-233/Product:** apoptosis regulator bcl-xs #status predicted <MA2>

Query Match	9.2%;	Score 99.5;	DB 2;	Length 233;
Best Local Similarity	23.1%;	Pred. No. 0.042;		
Matches 31;	Conservative 19;	Mismatches 57;	Indels 27;	Gaps 4

QY 44 PEAIV---RPSAARLRQIHRSFSSAYLYG---EGNFFELVALMAQSVLSDSGPMTGR 96  
D8 82 PMAAVQGLREAGDEPELRYRRAFSLTSGQIHITPGYVQSPQGVNMLPFD--GVNNGR 139  
QY 97 VYTLVTFAGTLLERGLVYTRMKMGFGQRLRQEGSDVAPRQOQRVALLSRRLMGCHRAW 156  
D8 140 IYAFSFGGLC-----VSSVDEKQVLVSRNAAMVATYLLDHLPEW 181

```
QY      157 LQAQGGWDGFCHFF 170
          :| ||| |
Db      182 IQENGWDTFVELY 195
```

## RESULT 13

```
bcl-x long - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
```

J. Immunol. 153, 4388-4398, 1994  
A:Title: Cloning and molecular characterization of mouse bcl-x in B and T lymphocytes  
;Reference number: 145055; MUID:95052604; PMID:796317

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-233 <RES>  
A:Cross-references: UNIPROT:Q64373, EMBL:U10101, NID:G506647, PIDD:AAA82173.1, PID:G506647  
A:Names: H. Michael, G.Y.; Takatsu, K.; Okuna, M.

A;Description: IL-5 inhibits anti-IgM-induced apoptosis in an immature B cell line throu  
A;Reference number: S52866

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-233 <KAM>  
A:Cross-references: EMBL:X83574; NID:695622; PIDD:CA58557.1; PID:695623  
A:Superfamily: bcl1 apoptosis regulator, inhibitory type

Query Match	9.2%;	Score 99.5;	DB 2;	Length 233;
Best Local Similarity	24.6%;	Pred. No. 0.042;		
Matches 34;	Conservative 17;	Mismatches 52;	Indels 35;	Gaps 6

QY	44	PEAAV---	LESAARLRQRIHSFFSYSLY---	-GNNPEVALLMADSVLDSGPTWGR	96
		----			
Db	82	PMAAVQALREAGDEPELRYRRAFSQLTSG	LIHTGTVAQSPQEVNLSFRD	-GVNNGR	139
QY	97	VYLVIVFAGTILERGEVLTVARMKKMGQERLAKGEGDVARDQGRVLVALLSRL---	MGQH	15	
Db	140	IYAFSFGGALCYE-----	SYDKKQIVVSRISIAEMVATYIADH	177	

```

QY      154 -RAWLQAGGWDGFCHFF 170
          | : | | | |
Db      178 LEPWIQENGWDTFVDLY 195

```

## RESULT 14

BCL-X protein - rat  
C;Species: Rattus norvegicus (Norway rat)

R;Michaelidis, T.M.  
submitted to the EMBL Data Library, November 1994  
A;Reference number: S51761

A:Accession: S51761  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-233 <MIC>  
A:Cross-references: UNIPROT:P53563; EMBL:X82537; NID:g607176; PIDD:CA557886.1; PID:g607176  
A:Experimental source: embryonic; brain  
A:Accession: S51762  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-125,189-233 <MI2>  
A:Cross-references: EMBL:X82537; NID:g607176; PIDD:CA557887.1; PID:g607178  
A:Experimental source: embryonic; brain  
A>Note: smaller form due to splicing  
C:Genetics:  
A:introns: 125/3  
C:Superfamily: bcl apoptosis regulator, inhibitory type

Query Match	9.2%	Score 99.5;	DB 2;	Length 233;
Best Local Similarity	24.6%;	Pred. No. 0.042;		
Matches 34;	Conservative 17;	Mismatches 52;	Indels 35;	Gaps 6

```

OY 44 PRAAV---LRSAARLRQIRHSFFSAYIGY-----PQNNFELVALMAADSVLSDSDGPPWGR 96
      : : : : :
Db 82 PMAAVQALREAGDEFELELYRRAFSDLTSLQILITFGTAYQSEFQVAVLELFRD--GVNWGR 139
      : : : : :
OY 97 VVTLVLTPEAGTILERGLVYARKKMGQPRLKEGDVAPRQGLVALLSRL---MGQH 153
      : : : : :
Db 140 IVAFPSGALCYE-----SYDKEMQVIVSRIRASMAITYLNDH 177

```

```

OY      154 -RAWLOAGGWDGFCCHF 170
          | : | | | | |
Db      178 LEPWIQENGWDTFVDLY 195

```

## RESULT 13

153744 gene bcl-2 protein - rat  
C\_Species: Rattus norvegicus (Norway rat)  
C\_Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
C\_Accession: 153744  
R\_Sato, T.; Irie, S.; Krajewski, S.; Reed, J.C.

**A;Title: Cloning and sequencing of a cDNA encoding the rat Bcl-2 protein**

A:Reference: NUMBER: 153744; MOLD:943503; ENID:018704  
A:Accession: 153744  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-36 <R>  
A:Cross-references: UNIPROT:P49950; GI:114680; NID:G408946; PID:AAA53662.1; PID:G408947  
C:Genetics:  
A:Gene: bcl-2

Query Match	9.2%;	Score 99.5;	DB 2;	Length 236;
Best Local Similarity	25.8%;	Pred. No. 0.043;		
Matches 40;	Conservative 17;	Mismatches 61;	Indels 37;	Gaps 7

QY 50 RSAARLRIOHRSFEFAYLGYPGNPELVYALMADSVLSDSPGPMGRVTLVTAGTILE 109  
Db 107 RDPAREMSQHLITPFA-----RGRF--ATYVELFPD--GVNMGRIVAFFEGGVMC- 155  
QY 110 RGPLVTARMKMGFQPLREKRGDVARDCQRLY--ALSSRLMGCH-RANLQAQSGMDG 165  
Db 156 --VGSVNREKSPSLVDNIATMRETYINRHLHTWIQNGGMDA 194

```

Qy      166 FCHERTPF-DLAFWRKQLVQAFSLCILTTPAIYL 199
      :  :  :  :  :  :  :  :  :  :  :  :
Db      195 FVELYGPMSRPLDFPSMLSKTLTSLALVGACTTL 229

```

Search completed: June 6, 2005, 10:44:23  
Job time : 40 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 06:21:30 ; Search time 282.584 Seconds  
(without alignment)  
5136.091 Million cell updates/sec

Title: US-10-071-174a-1

Perfect score: 887  
Sequence: 1 cggcgcaagaacacagcga.....ctctccttgagtgaaagaa 887

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55.6	6.3	2508	US-09-949-016-1466	Sequence 1466, Ap
2	55.6	6.3	97989	US-09-949-016-13208	Sequence 13208, A
3	55.2	6.2	5808	US-09-902-540-794	Sequence 794, App
4	54.8	6.2	6599	US-09-902-540-724	Sequence 724, App
5	54.4	6.1	1614	US-09-902-540-8389	Sequence 8389, Ap
6	54.4	6.1	7186	US-09-902-540-879	Sequence 879, App
7	52.8	6.0	1467	US-09-902-540-8509	Sequence 8509, Ap
8	52.8	6.0	6482	US-09-902-540-896	Sequence 896, App
9	52	5.9	3198	US-09-902-540-2522	Sequence 2522, Ap
10	52	5.9	17503	US-09-902-540-1114	Sequence 1114, Ap
11	51.4	5.8	967	US-09-902-540-7886	Sequence 7886, Ap
12	50.2	5.7	888	US-09-902-540-9238	Sequence 9238, Ap
13	50.2	5.7	2199	US-09-902-540-6388	Sequence 6388, Ap
14	50.2	5.7	2585	US-09-902-540-424	Sequence 424, App
15	50.2	5.7	9818	US-09-902-540-987	Sequence 987, App
16	49.8	5.6	1194	US-09-724-797-25	Sequence 25, Appl
17	49	5.5	678	US-09-902-540-4171	Sequence 4171, Ap
18	49	5.5	22301	US-09-902-540-1208	Sequence 1208, Ap
19	48.6	5.5	1702	US-09-902-540-8057	Sequence 8057, Ap
20	48.6	5.5	7160	US-09-902-540-821	Sequence 821, App
21	48.4	5.5	1092	US-09-252-991A-8481	Sequence 8481, Ap
22	48.4	5.5	3678	US-09-252-991A-8447	Sequence 8447, Ap
23	48	5.4	2561	US-09-616-289-48	Sequence 48, Appl
24	47.8	5.4	1362	US-09-724-797-31	Sequence 31, Appl
25	47.8	5.4	1365	US-09-252-991A-9680	Sequence 9680, Ap
26	47.8	5.4	1395	US-09-252-991A-9641	Sequence 9641, Ap
27	47.8	5.4	1443	US-09-252-991A-9615	Sequence 9615, Ap

28	47.6	5.4	2028	US-09-902-540-7619	Sequence 7619, Ap
29	47.6	5.4	7850	US-09-902-540-749	Sequence 749, App
30	47.4	5.3	1382	US-09-016-434-1256	Sequence 1256, Ap
31	47.4	5.3	1707	US-09-252-991A-10297	Sequence 10297, A
32	47.4	5.3	2511	US-09-252-991A-10560	Sequence 10560, A
33	47	5.3	1428	US-09-252-991A-14802	Sequence 14802, A
34	47	5.3	1668	US-09-252-991A-14950	Sequence 14950, A
35	46.8	5.3	711	US-09-902-540-9612	Sequence 9612, Ap
36	46.8	5.3	1380	US-09-902-540-5345	Sequence 5345, Ap
37	46.8	5.3	1393	US-09-902-540-9055	Sequence 9055, Ap
38	46.8	5.3	1633	US-09-902-540-6509	Sequence 6509, Ap
39	46.8	5.3	2605	US-09-902-540-470	Sequence 470, App
40	46.8	5.3	13299	US-09-902-540-968	Sequence 968, App
41	46.8	5.3	13631	US-09-902-540-1092	Sequence 1092, Ap
42	46.8	5.3	34662	US-09-902-540-1261	Sequence 1261, Ap
43	46.6	5.3	2403	US-09-949-016-1543	Sequence 1543, Ap
44	46.6	5.3	3252	US-10-327-189-3	Sequence 3, Appl
45	46.4	5.2	1344	US-09-252-991A-11619	Sequence 11619, A

## ALIGNMENTS

```
RESULT 1
US-09-949-016-1466
; Sequence 1466, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1466
; LENGTH: 2508
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1466

Query Match      6.3%  Score 55.6;  DB 4;  Length 2508;
Best Local Similarity 46.2%  Pred. No. 0.00036;
Matches 257;  Conservative 0;  Mismatches 294;  Indels 5;  Gaps 2;

QY      15 CAGCGAAGCCCGCCCGCCAGAGGCGGACCATGTTGACCACTTGCGGAGCGCA 74
DB      1624 CGGCAACCCCGCCCGCCCGCCAGAGGCGGACCATGTTGCGGAGCGG 1683
QY      75 CCACCATGCGGACCGCGGCGGAGCGGACCGGACTTGTGCGGACCACTTGCGG 134
DB      1684 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1743
QY      135 ACTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 193
DB      1744 ACTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1803
QY      194 CTGCGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 253
DB      1804 GTGCGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1863
QY      254 GCGTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 313
DB      1864 AG----CCCGGGAACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1919
QY      314 AGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 373
```

[illegible]

## RESULT 2

```

US-09-949-016-13208
; Sequence 13208, Application US/09949016
; Patent No. 6812339
;
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
;
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13208
;
; LENGTH: 97989
;
; TYPE: DNA
;
; ORGANISM: Human
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(97989)
;
; OTHER INFORMATION: n = A,T,C or G
;
US-09-949-016-13208

```

Query Match	6.3% ;	Score 55.6 ;	DB 4 ;	Length 97989 ;
Best Local Similarity	46.2% ;	Pred. No. 0.014 ;		
Matches 257 ;	Conservative 0 ;	Mismatches 294 ;	Indels 5 ;	Gaps 2 ;
QY	15	CAGCGAAGGCCCGGCCCCCAGCAGAGGCGCGACCAATGGTTGACCATTTGCGGAGCGCA	74	
Db	89050	CGGCGACCCCGCGCGCCCGCCCGCCCGCCAGGAACAAGTTCGCCCATCTTCGAGAGCAG	89109	
QY	75	CCACCATGAGCGCAGACCCGCTGCGGAGAGCGACCGAGCTGTTGTGCGCGCATCTTGGGAT	134	
Db	89110	TGTCGGTGGAGAGAGAGAGACGGGGGGGGGAGCAATCAAGATGCTCTTGAGAGGGCCGGGCGCGGG	89169	
QY	135	ACTGCGGCCCGGGAACCCGGCAACCCCGAGCGCG-CGCGCATCCAGCGCCGAGGCGCGCTGG	193	
Db	89170	ACTGCGGCCCGCAGCAGCTGGGGGGGAGAGAGCAGGCGCGGGGCTTCAGGGGTGTCTCCGACGG	89229	
QY	194	CTGCGCTCCGCGCGCGCCGACGATTACGGCAGATTACCGGCTCTTTTCTCCGCTTAACTC	253	
Db	89230	GTGCGCTTGCGCGCGCGCTGCTCTGTGTGATGATCTTCCGACACCGAGGTTTGAGGCCGGAAGC	89289	
QY	254	GGCTACCCCGGGGAACCGCTTCGAGCTGTGTGTGGCGCTGATGGCGGATTCGTCGCTCTCCAG	313	
Db	89290	AG-----CCCGAGAACCCCAAGCGCGCGGAGACCGCGTGAACGAGTCAACCCCGCTGTTCAG	89345	

QY	314	AGCCCCGGCCCA	CTGGGGCAAGTGTGA	CGCTGTGAC	CTTCGCAAGGAC	GTGCTG	373		
Db	89346	TGAGCTTCCCA <td>CCACCGGGC<td>CHAGCA</td><td>CAACGCCCCG<td>GAGGCTCCAGGCTGCTTGAGCTGG</td><td>89405</td></td></td>	CCACCGGGC <td>CHAGCA</td> <td>CAACGCCCCG<td>GAGGCTCCAGGCTGCTTGAGCTGG</td><td>89405</td></td>	CHAGCA	CAACGCCCCG <td>GAGGCTCCAGGCTGCTTGAGCTGG</td> <td>89405</td>	GAGGCTCCAGGCTGCTTGAGCTGG	89405		
QY	374	GAGAGAGGCGCT	GTGTGAC	CCGCGCGGTG <td>AGAGAGTGGGCGCTTCCAG</td> <td>CGCGCTTAAG</td> <td>433</td>	AGAGAGTGGGCGCTTCCAG	CGCGCTTAAG	433		
Db	89406	CAGAGAGAGAG	AGCTGAGCA	CTTCCCGAGT <td>GAGATTAAC</td> <td>CCCTTGTGAGATGAGACG</td> <td>89465</td>	GAGATTAAC	CCCTTGTGAGATGAGACG	89465		
QY	434	GAGAGAGAGG	CGACCTG	CGCGGAG	CTGCACAG	CGCTGTGAGCTG	CGCG	493	
Db	89466	CAITCCAGAAA	GACGCTTAC	AGCAAGGCG	CGCGCTG	CGCGGGCTTGGGGG	AGCAACAC	89525	
QY	494	CTCATGGG	CAGAC	CGCGCTG	GGCTG	CAGGCTC	AGGCTCAAGGCGCGG	CTTGTGTAC	553
Db	89526	CGCCCCAGGA	CCGCGGCTT	CGGGGTG	CAGAGACTG	CGGTCC	AGGCCCCAGAA	CAAGAG	89585
QY	554	TTCTTCAG	ACCCCT	569					
Db	89586	AACTTCAG	CGCGCT	89601					

### RESULT 3

```

US-09-902-540-794
Sequence 794, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 794
LENGTH: 5808
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-794

```

Query Match	5.2#:	Score	5.2:	DB 4:	Length	5808:	
Best Local Similarity	47.6#:	Pred.	No. 0.0061:				
Matches	162:	Conservative	0:	Mismatches	178:	Indels	0:
							Gaps
Qy	52	GATTGACCAATTGGCGGAGCGCACACCATGAGCCGACCCGCTGCGGAGCCGACCGAGCT	111				
Db	4805	GTTGACCATCTCCAGTCCCAAGAAAGATTGGCCGATGCGGCTGAGACCGGCCATGAT	486				
Qy	112	GTTGCTGCGCCACTTAACCTTGGGGTACTGTGCGCCCGGAAACCGGACATCCCGAGCCGCGCC	171				
Db	4865	GGAAGCCGACACACACGACGCGAGGTTCGCGGGGAGCATTTTTCACCAACGCGGATTCGCGC	492				
Qy	172	ATTCAGCCCGGAGGCCGCGCGTGGCTGCGCGCTCCGCGGCGCCAGAGTTACGGGACATTCACG	231				
Db	4925	CCATTCGCGGCGCTGTGTGGGGGGATGTTCTGGACGCGCCCGCGCTGTGTCGCGCGC	498				
Qy	232	GTCCCTTTTCTCCGCTACTCTCGGCTACCCCGGAAACCGCTTTCAGAGCTGTGGCCGAT	291				
Db	4985	CGCGGGCGCTGTGGGCTGTGCGAGGAAACCCGCGTGAACATACGGCCGTGTGTGCTGGG	504				
Qy	292	GGCGGATTCGCTGCTCTTCGACAGCGCCCGGCCCACTGGGGGCAAGTGTGAAGCTGT	351				
Db	5045	CGCGGTGCTCGCGCTGGGTGCTGTGTCTCAACCGCGGCCCGCGAAGGGAATGCGCAAGT	510				
Qy	352	GACCTTCGCAAGGACGTGCTGAGAGAGAGGGCCCGCTGGT	391				
Db	5105	GGGCGTGTGTGTGAACGTCTGGCCACCGGGGTGTGCTGGCG	514				

## RESULT 4

5105 GGGCGTGTGTGCACTGCTGGCCACGGGTGCTGCCG 5144

US-09-902-540-724  
; Sequence 724, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 724  
; LENGTH: 6599  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-724

Query Match 6.2%; Score 54.8; DB 4; Length 6599;  
Best Local Similarity 45.4%; Pred. No. 0.0081;  
Matches 197; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

QY 22 GGGCCGCGCCCGGAGGCGGAGCCGATGTTGACGAGTTGGGGAGCGGACCCAGCAT 81  
DB 2463 GAACTGCGCCGAGCACTGTATCTATCTCCGCTCGACCGGCTGATGATCCACAC 2522  
QY 82 GCGCGACCCGCTGCGGAGCGGACCGAGCTGTGCTGGCGGACGACTACTGAGGTA 141  
DB 2523 GCTGTACCGGACGACTGTGCGCGCGGCTGACGCGGCTGACGCGGCTACTACTA 2582  
QY 142 CCGGAGACCCGACCCCGGAGCGGCGGATCCAGCGGCGGCGGCTGCGCTC 201  
DB 2583 CTGTGTCAGCGGCTCACTGCTGCTGCGGAGATGAGCAACCGGCTTCTCTC 2642  
QY 202 GCGCGCGCGGAGTTAGGAGATTCACCGGCTTTTTCGCGCTACCTCGGTA 261  
DB 2643 GCGCTAGTGTGGGCTGCGGCTGAGCTCGGATTCGCTGCGGAGCTGGGCGG 2702  
QY 262 GCGGACCGCTTTCAGCTGTGTGCGGCTGATGCGGAGTTCCGCTCTCGACAGCCCG 321  
DB 2703 CACCTGCGCGCGGCTGCTGCGGCTGCTGCTGCGGCTGCTGCGGCTGCTG 2762  
QY 322 CCGGACCTGCGGAGAGTGTGAGCGCTGTGACCTTTCGAGGAGCGCTGAGAGG 381  
DB 2763 CATCACTAGTGTGCGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 2822  
QY 382 GCGGCTGTGACCGCGGCTGAGAGATGTGGGCTTTCAGCGGCGGCTGAGAGAG 441  
DB 2823 GCTGAGGACACGCGCTGCGGCGGAGCGTGGGCGTGGGCGTGGCTGCTGCTG 2882  
QY 442 GGGGACGCTGCGG 455  
DB 2883 GGTGCTCTCTTCC 2896

RESULT 5  
US-09-902-540-8389  
; Sequence 8389, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10

NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 8389  
; LENGTH: 1614  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-8389

Query Match 6.1%; Score 54.4; DB 4; Length 1614;  
Best Local Similarity 46.5%; Pred. No. 0.0061;  
Matches 175; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 130 GGGGTACTGCGCCCGGAGAACCGGACCCCGAGCGGCGGATTCACGCGGAGCGC 189  
DB 645 GGAATAGTGGGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 704  
QY 190 GGTGCTGCTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 249  
DB 705 CTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 764  
QY 250 CCGGCTGACCGCGGAGAACCGGCTTGAAGTGTGCGGCTGATGAGGAGGAGGAGGAGGAG 309  
DB 765 GCGCTTTCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 824  
QY 310 GAGAGCGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 369  
DB 825 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 884  
QY 370 GCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 429  
DB 885 CCGGCTGCTTTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 944  
QY 430 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489  
DB 945 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1004  
QY 490 GCGGCTGATGCGGCGGAG 505  
DB 1005 CCGGCTGATGCGGCGGAG 1020

RESULT 6  
US-09-902-540-879/c  
; Sequence 879, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 879  
; LENGTH: 7186  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-879

Query Match 6.1%; Score 54.4; DB 4; Length 7186;  
Best Local Similarity 46.5%; Pred. No. 0.0011;  
Matches 175; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 130 GGGGTACTGCGCCCGGAGAACCGGACCCCGAGCGGCGGATTCACGCGGAGCGC 189  
DB 5491 GGAATAGTGGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5432  
QY 190 GGTGCTGCTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 249  
DB 5431 CTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5372





Query Match 5.9%; Score 52; DB 4; Length 3198;  
Best Local Similarity 53.4%; Pred. No. 0.0032;  
Matches 109; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 18 CGAAGCCCCGCCCCCGAGAGGCGGACCATGTTGACGTTGGGGAGCGCACCA 77  
DB 764 CGAAGACCATCTCGGTCTACAGAGAGCTGGCCAAATTCATCAGAGAGCGGCGCGTCA 823  
QY 78 CCATGCGCCGACCCCGCTGGGGAGCGCACCGAGCTTGTCTGGCCCACTACTGGGGTACT 137  
DB 824 CCGAGGCGGAGGCTGTCTGACGCGAATTTGAAGTCTGAGCCCGAGAGACCCGAGACTGG 883  
QY 138 GCGCCCGGGAACCGCGGACACCCCGAGCGGCGGCATCCACGCGGAGCGCGGTGCTGC 197  
DB 884 TCGGCGCGCGCGCGCGCGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCACTCAGC 943  
QY 198 GCTCCGCGCGCGCGCGCGGTTACGCG 221  
DB 944 CCGCCCGCGCGCGCGCGCGCGCGCGC 967

RESULT 10  
US-09-902-540-1114/c  
; Sequence 1114, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(115849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1114  
; LENGTH: 17503  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-1114

Query Match 5.9%; Score 52; DB 4; Length 17503;  
Best Local Similarity 53.4%; Pred. No. 0.0059;  
Matches 109; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 18 CGAAGCCCCGCCCCCGAGAGGCGGACCATGTTGACGTTGGGGAGCGCACCA 77  
DB 10937 CGAAGACCATCTCGGTCTACAGAGAGCTGGCCAAATTCATCAGAGAGCGGCGCGTCA 10878  
QY 78 CCATGCGCCGACCCCGCTGGGGAGCGCACCGAGCTTGTCTGGCCCACTACTGGGGTACT 137  
DB 10877 CCGAGGCGGAGGCTGTCTGACGCGAATTTGAAGTCTGAGCCCGAGAGACCCGAGACTGG 10818  
QY 138 GCGCCCGGGAACCGCGGACACCCCGAGCGGCGGCATCCACGCGGAGCGCGGTGCTGC 197  
DB 10817 TCGGCGCGCGCGCGCGCGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCACTCAGC 10758  
QY 198 GCTCCGCGCGCGCGCGGTTACGCG 221  
DB 10757 CCGCCCGCGCGCGCGCGCGCGCGCGC 10734

RESULT 11  
US-09-902-540-7886  
; Sequence 7886, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(115849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 7886  
; LENGTH: 967  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-7886

Query Match 5.8%; Score 51.4; DB 4; Length 967;  
Best Local Similarity 48.2%; Pred. No. 0.0029;  
Matches 145; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 91 GCTGCGGAGCGCACCGAGCTGTGCTGCGCCGACTACTCTGGGGTACTGCGCGGGAAC 150  
DB 3 GCGGCTGGAACCGGCGCATGTGTGAGGCGGAGAGAGAGCGCACCGGTGTTCGCGGGGACT 62  
QY 151 CGGACCCCCGAGCGCGGCGCATCCAGCGCCGAGGCGCGGCTGCGCGCGCGC 210  
DB 63 TGTACACGCGCGGCTGTCCGCAATCCGCGCGCTGCGGCGGCGGTGCTGCGGAGCGC 122  
QY 211 CAGGTTAGGCAATACCGGCTTTTCTCCGCGCTACTCGGCGTACCCCGGGAACCG 270  
DB 123 GCGGCGCTGTTCGTCGCGCGCGCGCGCGCTGTGCGCTGCGAGAGACCCCGCGTGAAC 182  
QY 271 CTTCGAGCTGTGTGCGCTGATGAGCGGATTCGCTGCTCCGACAGCCCCGCGCACTGG 330  
DB 183 ATAGCGCGCTGTGTGCGCTGCGCGCGCGCGCGCTGCGCGCTGTGCGTCAACCGCGCGC 242  
QY 331 GGGCAGAGTGTGACGCTGCTGACCTTGACGAGGAGCGTGTGAGAGAGGCGCGTGT 390  
DB 243 CGGAGAGGAGATGAGCGAGGTGCGCGCGCTGTGCTGCACTGCTGCGCACGCGGCTGGC 302  
QY 391 G 391  
DB 303 G 303

RESULT 12  
US-09-902-540-9238  
; Sequence 9238, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(115849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 9238  
; LENGTH: 888  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-9238

Query Match 5.7%; Score 50.2; DB 4; Length 888;  
Best Local Similarity 46.2%; Pred. No. 0.0056;  
Matches 166; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 17 GCGAAGCGCGCGCGCGCGCGAGAGCGCGGACCATGTTGACGTTGGGGAGCGCACCC 76  
DB 13 GCGCTGAGGCTGCGCTGTGGGGAGAGGTGTCGCGCGGTGTGCTGAGCGAGCGGAG 72  
QY 77 ACATGCGGACCGGCTGCGGAGAGCGCACGAGCTGTGCTGCGGCACTACTCTGGGGTAC 136

Db 73 ACGCCCGCGAGGTCCTGAGAGTGGCGACGCGAGGGGTGAGAGACTGCTGTGTGCG 132  
QY 137 TGGCGCCGGGAAACCCGGGACACCCCGAGCGCGCCATCCAGCCCGGCGCGTGTGCG 196  
Db 133 CCCCTGCGACGACGAGCTGGCGCGCGGTTGGGAGTGTGGCGGGTGGCGCGTCCCG 192  
QY 197 CGCTCCGCGCGCGCGAGGTTACGCGAGATTCAACGGTCTTTTCTCCGCTACTCCGCG 256  
Db 193 GCGCGGTGTGTGGCGGTGATACCGTGTGCGTTGCGCCCTGCTGCTGTGGCGCGCGCT 252  
QY 257 TACCCCGGGAACCGCTTGAAGTGTGGCGCTGATGGCGGATTCCGTCTCCGACAGC 316  
Db 253 GCGCCCGTGGCGTGGAGCGCGGCTGGCGCTGCTGAGAGCGGTGTGATCATGCTCTC 312  
QY 317 CCGCGCCCGACCTGGGGGAGAGTGTGACGCTCTGAGACTTTCGACGGAGAGCTGTGTA 375  
Db 313 TACAGCGCCACCGTGGGCGGTGGCGCGCGCGCTGGCCGAACAGGGGGAAGCGCCGA 371

## RESULT 13

US-09-902-540-6388  
; Sequence 6388, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OR INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 6388  
; LENGTH: 2199  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-6388

Query Match 5.7%; Score 50.2; DB 4; Length 2199;  
Best Local Similarity 44.6%; Pred. No. 0.0079;  
Matches 196; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

QY 91 GCTGCGGAGCGCACCGAGCTTTCCTGCGCGACTACTCTGGGTACTGCGCGCGGAAAC 150  
Db 378 GCTGTGTGTGGGCAAGCGCGCGCGCGAGGCGCATCTCCGGGCGCGCGGTGTGAC 437  
QY 151 CGGCAACCCCGAGCGCGCGCATTCACGCGCGGAGCGCGCGTGTGCGCTCCGCGCGC 210  
Db 438 CGGCACTTTCGCGGT 497  
QY 211 CAGGTTACGGCAGATTCAACCGGTCTTTTCTCCGCTACTCTGCGCTAACCCCGGAAAC 270  
Db 498 CTTCGTGCGGTGAGAGCGCGGACGACACTGTACGGGCTGCGCGGAGCGCGCGGAA 557  
QY 271 CTTCGAGCTGT 330  
Db 558 GTCCCGCGCTGT 617  
QY 331 GGGCAGAGTGTGAGCGCTGTGACCTTTCGACGAGAGCGCTGTGAGAGAGGCGCTGTGT 390  
Db 618 GCTGT 677  
QY 391 GACCGCCCGGTGAAAGAAAGTGGGCTTTCAGCGCGGCTAAAGAGAGAGAGGAGCGT 450  
Db 678 CAATCTCGGAGGAGGCGGTGCGCGGATGTGCTCCAGATGCGGCGAGGCTCCGGCTACG 737  
QY 451 CGCGCGGAGACTGACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510  
Db 738 CGACCGCGGCACTTCTTCTGT 797

QY 511 CGCTGTGCTGAGGCTCAG 529  
Db 798 GCGCTGT 816

## RESULT 14

US-09-902-540-424  
; Sequence 424, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OR INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 424  
; LENGTH: 2585  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-424

Query Match 5.7%; Score 50.2; DB 4; Length 2585;  
Best Local Similarity 44.6%; Pred. No. 0.0083;  
Matches 196; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

QY 91 GCTGCGGAGCGCACCGAGCTTTCCTGCGCGACTACTCTGGGTACTGCGCGCGGAAAC 150  
Db 716 GCTGTGTGTGGGCAAGCGCGCGCGCGAGGCGCATCTCCGGGCGCGCGGTGTGAC 775  
QY 151 CGGCAACCCCGAGCGCGCGCATTCACGCGCGGAGCGCGCGTGTGCGCTCCGCGCGC 210  
Db 776 CGGCACTTTCGCGGT 835  
QY 211 CAGGTTACGGCAGATTCAACCGGTCTTTTCTCCGCTACTCTGCGCTAACCCCGGAAAC 270  
Db 836 CTTCGTGCGGTGAGAGCGCGGACGACACTGTACGGGCTGTGCGGAGCGCGCGGAA 895  
QY 271 CTTCGAGCTGT 330  
Db 896 GTCCCGCGCTGT 955  
QY 331 GGGCAGAGTGTGAGCGCTGTGACCTTTCGACGAGAGCGCTGTGAGAGAGGCGCTGTGT 390  
Db 956 GCTGT 1015  
QY 391 GACCGCCCGGTGAAAGAAAGTGGGCTTTCAGCGCGGCTAAAGAGAGAGAGGAGCGT 450  
Db 1016 CAATCTCGGAGGAGCGGTGCGCGGATGTGCTCCAGATGCGGCGAGACTCCGCTACG 1075  
QY 451 CGCGCGGAGACTGACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510  
Db 1076 CGACCGGCACTTCTTCTGT 1135  
QY 511 CGCTGTGTGAGGCTCAG 529  
Db 1136 GCGCTGT 1154

## RESULT 15

US-09-902-540-987  
; Sequence 987, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.



**This Page Blank (uspto)**

```

Query Match 100.0%; Score 887; DB 16; Length 887;
Best local Similarity 100.0%; Pred. No. 3.4e-256;
Matches 887; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

```
QY 61 GTTGGGAGGCGCACACCATATGCGACCCGCTGCGGAGCGACCGAGCTGTGCTGAC 120
Db 61 GTTGGGAGGCGCACACCATATGCGACCCGCTGCGGAGCGACCGAGCTGTGCTGAC 120
QY 121 CGACTACTGAGGAGTACTGCGCCCGGAAACCCCGGACCCCGAGCCGGCCATCCACGCC 180
Db 121 CGACTACTGAGGAGTACTGCGCCCGGAAACCCCGGACCCCGAGCCGGCCATCCACGCC 180
QY 181 CGAGGCGCGGCTGCTGCGGCTCCGCGGCGCGCAGGTTACGGGAGATTCACCCGCTCTTTT 240
Db 181 CGAGGCGCGGCTGCTGCGGCTCCGCGGCGCGCAGGTTACGGGAGATTCACCCGCTCTTTT 240
QY 241 CTCGCGCTACCTGCGCTACCCCGGAAACCGCTTGAAGTGTGCGCTGAGTGGCGGATTC 300
Db 241 CTCGCGCTACCTGCGCTACCCCGGAAACCGCTTGAAGTGTGCGCTGAGTGGCGGATTC 300
QY 301 CGTGTCTTCCGACACGCCCGGCCCACTGCGGCGGAGAGTGTGAGGCTGTGACCTTGGC 360
Db 301 CGTGTCTTCCGACACGCCCGGCCCACTGCGGCGGAGAGTGTGAGGCTGTGACCTTGGC 360
QY 361 AGGAGCGCTGCTGAGAGAGGCGCGCTGCTGACCGCGCGGTGAGAGAGAGTGGGCTTCCA 420
Db 361 AGGAGCGCTGCTGAGAGAGGCGCGCTGCTGACCGCGCGGTGAGAGAGAGTGGGCTTCCA 420
QY 421 GCGCGGCTTAAAGAGAGAGAGGCGAGCTGCGCGCGGAACTGCGCAGCGCTGTGCGCTT 480
Db 421 GCGCGGCTTAAAGAGAGAGAGGCGAGCTGCGCGCGGAACTGCGCAGCGCTGTGCGCTT 480
QY 481 GCTGAGCTTCCGCGCTCAATGCGGCGACACCGCGCTGCGGAGGCTCAAGGCGGCTGCGA 540
Db 481 GCTGAGCTTCCGCGCTCAATGCGGCGACACCGCGCTGCGGAGGCTCAAGGCGGCTGCGA 540
QY 541 TGGCTTTTGTCACTCTTCAAGGACCCCTTTTCACTGCGCTTTTGAAGAAACAGCTGCT 600
Db 541 TGGCTTTTGTCACTCTTCAAGGACCCCTTTTCACTGCGCTTTTGAAGAAACAGCTGCT 600
QY 601 CCAAGCTTTTCTGTATGCTTGTAAACAAGCCTTCACTTATCTGTGACAGATTAAT 660
Db 601 CCAAGCTTTTCTGTATGCTTGTAAACAAGCCTTCACTTATCTGTGACAGATTAAT 660
QY 661 ATGACTTTTAAACCTTTTAAACCCGCTTCTCACTGCGGCACTGTGACCACTTAAATGACAG 720
Db 661 ATGACTTTTAAACCTTTTAAACCCGCTTCTCACTGCGGCACTGTGACCACTTAAATGACAG 720
QY 721 ATGTGTGAGAACAAAGACTGAGGAGAAAGCACTTCCCAAGGCTTTTATCTGA 780
Db 721 ATGTGTGAGAACAAAGACTGAGGAGAAAGCACTTCCCAAGGCTTTTATCTGA 780
QY 781 ATGCATTAACAAGAGTCTGAGGTGATTTGGCCAGATGTTTAACTTGTGACAGTACT 840
Db 781 ATGCATTAACAAGAGTCTGAGGTGATTTGGCCAGATGTTTAACTTGTGACAGTACT 840
QY 841 CAGGTGTGAGGACAAAGATGCAATGCGCTCTTCTTGAAGTGAAGA 887
Db 841 CAGGTGTGAGGACAAAGATGCAATGCGCTCTTCTTGAAGTGAAGA 887

RESULT 2
US-09-912-599-3
; Sequence 3, Application US/09912599
; Patent No. US20020106731A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Bcl-2-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO44P1
; CURRENT APPLICATION NUMBER: US/09/912,599
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: PCT/US01/03080
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/179,487
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/180,697
; PRIOR FILING DATE: 2000-02-07
```

```
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-912-599-3

Query Match      57.1%; Score 506.8; DB 9; Length 522;
Best Local Similarity 99.4%; Pred. No. 7.2e-142;
Matches 519; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 270 GCTTTCAGCTGTGTGCGCTGATGCGGATTCGCTGCTCCGACAGCCCGGCGCCACT 329
Db 1 GCTTTCAGCTGTGTGCGCTGATGCGGATTCGCTGCTCCGACAGCCCGGCGCCACT 60
QY 330 GGGGCAAGTGTGACGCTGCTGACCTTTCGAGGAGAGCGCTGCTGAGAGAGGCGCTGCG 389
Db 61 GGGGCAAGTGTGACGCTGCTGACCTTTCGAGGAGAGCGCTGCTGAGAGAGGCGCTGCG 120
QY 390 TGACCGCCCGGTGAGAGAGTGGGGCTTCAGCCGCGCTTAAAGAGAGAGGCGAGC 449
Db 121 TGACCGCCCGGTGAGAGAGTGGGGCTTCAGCCGCGCTTAAAGAGAGAGGCGAGC 180
QY 450 TCGCCCGGAGCTGCGCAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509
Db 181 TCGCCCGGAGCTGCGCAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 510 GCGCTGCTGCTGAGGCTTGAAGGCGGCTGAGATGCTTTGTCATCTTTCAGAGCCCT 569
Db 241 GCGCTGCTGCTGAGGCTTGAAGGCGGCTGAGATGCTTTGTCATCTTTCAGAGCCCT 300
QY 570 TTCACTGCTTTTGGAGAAACAGCTGTGCTGCTGCTTCTGTCATGCTTGTAAACA 629
Db 301 TTCACTGCTTTTGGAGAAACAGCTGTGCTGCTGCTTCTGTCATGCTTGTAAACA 360
QY 630 CAGCTTATTTATCTGTGACAGGATTTATGAGTTTAAACCTTTAAACCCGCTTCT 689
Db 361 CAGCTTATTTATCTGTGACAGGATTTATGAGTTTAAACCTTTAAACCCGCTTCT 420
QY 690 AACTGCGCAATGTGACCACTTAAATGACAGATGTGAGAAACAAGAGGAAAGC 749
Db 421 AACTGCGCAATGTGACCACTTAAATGACAGATGTGAGAAACAAGAGGAAAGC 480
QY 750 ACCTTCCCAAGCCAGAGC-TTTTATCTGATGATACAA 790
Db 481 ACCTTCCCAAGCCAGAGC-TTTTATCTGATGATACAA 522

RESULT 3
US-09-764-847-311
; Sequence 311, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 311
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,c,g, or c
; NAME/KEY: SITE
; LOCATION: (20)
; OTHER INFORMATION: n equals a,t,c,g, or c
```

NAME/KEY: SITE  
LOCATION: (26)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (28)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (53)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-847-311

Query Match 51.5%; Score 457.2; DB 9; Length 548;  
Best Local Similarity 98.2%; Pred. No. 6.2e-127;  
Matches 481; Conservative 2; Mismatches 5; Indels 2; Gaps 2;

QY 259 CCCCCGAAACCGCTTCGAGCTGTGTGGCGCTGTATGGCGGATTCGCTGCTCCGACAGCCC 318  
DB 61 CCAAGCGCTCGGCTTCGAGCTGTGTGGCGCTGTATGGCGGATTCGCTGCTCCGACAGCCC 120  
QY 319 CCGCCCCACCTGTGGGCAAGTGTGACGCTGTGACCTTCGACAGGAGCGCTGTGAGAG 378  
DB 121 CCGCCCCACCT-GGGCAAGTGTGACGCTGTGACCTTCGACAGGAGCGCTGTGAGAG 179  
QY 379 AGGGCCGCTGTGACCGCCCGGTGAAAGTGGGGCTTCAGCCGCGCTAAAGAGACA 438  
DB 180 AGGGCCGCTGTGACCGCCCGGTGAAAGTGGGGCTTCAGCCGCGCTAAAGAGACA 239  
QY 439 GGAAGGCGACGTGCGCCCGGAGACTGCAAGCGCTGTGCTGTGAGCTGCGGCTCAT 498  
DB 240 GGAAGGCGACGTGCGCCCGGAGACTGCAAGCGCTGTGCTGTGAGCTGCGGCTCAT 299  
QY 499 GGGGCAAGACCGCGCTGTGACGCTCAGGCGCGCTGGAGTGGCTTTGTGCACTTCTT 558  
DB 300 GGGGCAAGACCG- GCTGTGCTGACAGCTCAGGCGCGCTGGAGTGGCTTTGTGCACTTCTT 358  
QY 559 CAGGACCCCTTTTCACCTGCTGCTTTTGGAGAAAACAGCTGTGCAAGCTTTTCTGTATG 618  
DB 359 CAGGACCCCTTTTCACCTGCTGCTTTTGGAGAAAACAGCTGTGCAAGCTTTTCTGTATG 418  
QY 619 CTGTTAACAACAGCTTCATTTATCTGTGACAGATTAATTAAGATTAAACCTTTT 678  
DB 419 CTGTTAACAACAGCTTCATTTATCTGTGACAGATTAATTAAGATTAAACCTTTT 478  
QY 679 AACCGCTTCTACCTGCGCAACTGTGACCACTAAATGACAGATGTGAGAACAGAAC 738  
DB 479 AACCGCTTCTACCTGCGCAACTGTGACCACTAAATGACAGATGTGAGAACAGAAC 538  
QY 739 TGAGGGAAG 748  
DB 539 TGAGGGAAG 548

## RESULT 4

US-10-092-154-311  
Sequence 311, Application US/10092154  
Publication No. US20030054375A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC009C1  
CURRENT APPLICATION NUMBER: US/10/092,154  
CURRENT FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 2003  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 311  
LENGTH: 548  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (16)  
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature  
LOCATION: (20)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (26)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (28)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (53)  
OTHER INFORMATION: n equals a,t,g, or c  
US-10-092-154-311

Query Match 51.5%; Score 457.2; DB 14; Length 548;  
Best Local Similarity 98.2%; Pred. No. 6.2e-127;  
Matches 481; Conservative 2; Mismatches 5; Indels 2; Gaps 2;

QY 259 CCCCCGAAACCGCTTCGAGCTGTGTGGCGCTGTATGGCGGATTCGCTGCTCCGACAGCCC 318  
DB 61 CCAAGCGCTCGGCTTCGAGCTGTGTGGCGCTGTATGGCGGATTCGCTGCTCCGACAGCCC 120  
QY 319 CCGCCCCACCTGTGGGCAAGTGTGACGCTGTGACCTTCGACAGGAGCGCTGTGAGAG 378  
DB 121 CCGCCCCACCT-GGGCAAGTGTGACGCTGTGACCTTCGACAGGAGCGCTGTGAGAG 179  
QY 121 CCGCCCCACCT-GGGCAAGTGTGACGCTGTGACCTTCGACAGGAGCGCTGTGAGAG 179  
QY 379 AGGGCCGCTGTGACCGCCCGGTGAAAGTGGGGCTTCAGCCGCGCTAAAGAGACA 438  
DB 180 AGGGCCGCTGTGACCGCCCGGTGAAAGTGGGGCTTCAGCCGCGCTAAAGAGACA 239  
QY 439 GGAAGGCGACGTGCGCCCGGAGACTGCAAGCGCTGTGCTGTGAGCTGCGGCTCAT 498  
DB 240 GGAAGGCGACGTGCGCCCGGAGACTGCAAGCGCTGTGCTGTGAGCTGCGGCTCAT 299  
QY 499 GGGGCAAGACCGCGCTGTGACGCTCAGGCGCGCTGGAGTGGCTTTGTGCACTTCTT 558  
DB 300 GGGGCAAGACCG- GCTGTGCTGACAGCTCAGGCGCGCTGGAGTGGCTTTGTGCACTTCTT 358  
QY 559 CAGGACCCCTTTTCACCTGCTGCTTTTGGAGAAAACAGCTGTGCAAGCTTTTCTGTATG 618  
DB 359 CAGGACCCCTTTTCACCTGCTGCTTTTGGAGAAAACAGCTGTGCAAGCTTTTCTGTATG 418  
QY 619 CTGTTAACAACAGCTTCATTTATCTGTGACAGATTAATTAAGATTAAACCTTTT 678  
DB 419 CTGTTAACAACAGCTTCATTTATCTGTGACAGATTAATTAAGATTAAACCTTTT 478  
QY 679 AACCGCTTCTACCTGCGCAACTGTGACCACTAAATGACAGATGTGAGAACAGAAC 738  
DB 479 AACCGCTTCTACCTGCGCAACTGTGACCACTAAATGACAGATGTGAGAACAGAAC 538  
QY 739 TGAGGGAAG 748  
DB 539 TGAGGGAAG 548

## RESULT 5

US-10-363-345A-30995  
Sequence 30995, Application US/10363345A  
Publication No. US20040234960A1  
GENERAL INFORMATION:  
APPLICANT: Alexander Olek  
APPLICANT: Christian Piepenbrock  
TITLE OF INVENTION: Method for determining the degree of methylation of defined  
FILE REFERENCE: E01/1227  
CURRENT APPLICATION NUMBER: US/10/363,345A  
CURRENT FILING DATE: 2003-03-03  
NUMBER OF SEQ ID NOS: 40712  
SEQ ID NO 30995  
LENGTH: 874  
TYPE: DNA  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
OTHER INFORMATION: Cpg-island No: 30995  
US-10-363-345A-30995

Query Match 36.6%; Score 324.6; DB 20; Length 874;  
Best Local Similarity 75.1%; Pred. No. 5.9e-87;  
Matches 405; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```
OY 1 CGGGCCAGAAACACGAGAGGCGCCGCCCGGACGAGAGCCGACCATGTTGACCA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 100 CGGGTTAAGAAATTAAGGAAAGTTGCGTTTATTAAGTAGAGTCCGATTAATGATTA 159
OY 61 GTTGGGAGAGCGACACCAATAGCCGACCGCTGCGGAGCGACACGAGCTGTTGCTGC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 160 GTTGGGAGAGCGATTAATTAATGATGATTCGTTGCGGAGAGCGATACGATGTTGTTG 219
OY 121 CGACTACCTGGGGTACTGCGCCCGGAAACCGGACACCCCGAGCGCGCATCCAGCC 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 220 CGATTATTTGGGGTATGTCGTTCCGGAATTCGATATTTTCAGATCGCGATTATTA 279
OY 181 CGAGCGCGCGCTGCTGCGCTCCGCGGCGGCAAGTTAGCGAGATTCACCGTCTT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 280 CGAGGTGCTGCTGTTGCGGTTGCGGTTAGTTAGTTAGTTATTCGTTTCTT 339
OY 241 CTCGCTACCTGCGCTACCCCGGAGACCGCTTCGAGCTGTTGCGCTGATGCGGATTC 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 340 TTTGCTTATTTTCGTTATTTTCGGAATCGTTTCAGATTGGTGCGTTGATGCGGATTT 399
OY 301 CGTGTCTCCGACAGACCCCGGCCCACTGCGGAGAGTGTGACGCTGTCGCTGCGC 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 400 CGTGTCTCCGACAGAGTTGCGTTTATTTGAGGATGAGTGAAGTTGATTTTCT 459
OY 361 AGGACGCTGCTGAGAGAGAGGCGCTGTGACCGCCCGGTGGAAGTGGGCTTCCA 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 460 AGGACGTTGTTGAGAGAGGCTGTTGATGATCGTTGTTGAGGCTTGAAGAGTGGG 519
OY 421 GCCGCGCTAAAGAGAGAGAGGAGCGTCCCGGCACTGCGCGCTGTGAGCTT 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 520 GTCCGCGTTAAAGAGAGTGAAGGAGCGACGCTTGGGATTTGACGTTTGATGTTT 579
OY 481 GCTGAGCTCGCGCTCATATGGGAGACACCGCGCTGCTGACAGCTCAAGGCGGCTG 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 580 GTTGAATTCGCGGTTTATGAGGATGATTCGCTTTGTTGAGGCTTGAAGGCGGTT 638
```

## RESULT 6

US-10-363-345A-30996/c  
Sequence 30996, Application US/10363345A  
Publication No. US20040234960A1  
GENERAL INFORMATION:  
APPLICANT: Alexander Olek  
APPLICANT: Christian Piepenbrock  
TITLE OF INVENTION: Method for determining the degree of methylation of defined  
FILE REFERENCE: E01/1227  
CURRENT APPLICATION NUMBER: US/10/363,345A  
CURRENT FILING DATE: 2003-03-03  
NUMBER OF SEQ ID NOS: 40712  
SEQ ID NO 30996  
LENGTH: 874  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
OTHER INFORMATION: Cpg-island No: 30996  
US-10-363-345A-30996

Query Match 36.6%; Score 324.6; DB 20; Length 874;  
Best Local Similarity 75.1%; Pred. No. 5.9e-87;  
Matches 405; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```
OY 1 CGGGCCAGAAACACGAGAGGCGCCGCCCGGACGAGAGCCGACCATGTTGACCA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 775 CGGGTTAAGAAATTAAGGAAAGTTGCGTTTATTAAGTAGAGTCCGATTAATGATTA 716
OY 61 GTTGGGAGAGCGACACCAATAGCCGACCGCTGCGGAGCGACACGAGCTGTTGCTGC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 715 GTTGGGAGAGCGATTAATTAATGATGATTCGTTGCGGAGAGCGATACGATGTTGTTG 656
OY 121 CGACTACCTGGGGTACTGCGCCCGGAAACCGGACACCCCGAGCGCGCATCCAGCC 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 655 CGATTATTTGGGGTATGTCGTTCCGGAATTCGATATTTTCAGATCGCGATTATTA 596
OY 181 CGAGCGCGCGCTGCTGCGCTCCGCGGCGGCAAGTTAGCGAGATTCACCGTCTT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 595 CGAGGTGCTGCTGTTGCGGTTGCGGTTAGTTAGTTAGTTATTCGTTTCTT 536
OY 241 CTCGCTACCTGCGCTACCCCGGAGACCGTTTGAAGTGTGCGCTGATGCGGATTC 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 535 TTTGCTTATTTTCGTTATTTTCGGAATTCGATATTTTCAGATGCGGATTGCGGATTT 476
OY 301 CGTGTCTCCGACAGACCCCGGCCCACTGCGGAGAGTGTGACGCTGTCGCTGCGC 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 475 CGTGTCTCCGACAGAGTTGCGGTTTATTTGAGGATGAGTGAAGTTCGATTTCT 416
OY 361 AGGACGCTGCTGAGAGAGGCGCTGTGACCGCCCGGTGGAAGTGGGCTTCCA 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 415 AGGACGTTGTTGAGAGAGGCGCTGTGATGATGATGATGATGATGATGATGATGAT 356
OY 421 GCCGCGCTAAAGAGAGAGGCGACGTCGCCCGGACCTGCAAGGCGCTGTGCGCTT 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 355 GTCCGCGTTAAAGAGAGAGGCGACGTCGTTGCGGATTTGTAAGCGTTGCGGATTT 296
OY 481 GCTGAGCTCGCGCTCATATGGGAGACACCGCGCTGCTGACGCTCAAGGCGGCTG 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 295 GTTGAATTCGCGGTTTATGAGGATGATTCGCTTTGTTGAGGCTTGAAGGCGGTT 237
```

## RESULT 7

US-10-363-345A-30993/c  
Sequence 30993, Application US/10363345A  
Publication No. US20040234960A1  
GENERAL INFORMATION:  
APPLICANT: Alexander Olek  
APPLICANT: Christian Piepenbrock  
TITLE OF INVENTION: Method for determining the degree of methylation of defined  
FILE REFERENCE: E01/1227  
CURRENT APPLICATION NUMBER: US/10/363,345A  
CURRENT FILING DATE: 2003-03-03  
NUMBER OF SEQ ID NOS: 40712  
SEQ ID NO 30993  
LENGTH: 874  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
OTHER INFORMATION: Cpg-island No: 30993  
US-10-363-345A-30993

Query Match 36.1%; Score 320; DB 20; Length 874;  
Best Local Similarity 74.8%; Pred. No. 1.4e-85;  
Matches 401; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

```
OY 1 CGGGCCAGAAACACGAGAGGCGCCGCCCGGACGAGAGCCGACCATGTTGACCA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 775 CGAACCAAAACCAACCAAGAAACCCGACCCCAACCAAAACCAACCAATTAATTA 716
OY 61 GTTGGGAGAGCGACACCAATAGCCGACCGCTGCGGAGCGACACGAGCTGTTGCTGC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 715 ATTAGAAAGAGCGACACCAATAGCCGCTGAGAAAGAGCGACCGAATTAATTAAT 656
OY 121 CGACTACCTGGGGTACTGCGCCCGGAAACCGGACACCCCGAGCGCGCATTCACGCC 180
```



Db 655 CGACTACCTAAATACTACGCCCCGAAACCCGACACCCCGAACCGACGCAATTCACGCC 596  
Qy 181 CGAGGCGCGCGTGTGCTGCTCCGCGCGCCGCAAGTTACGAGATTACCGGTCTTTT 240  
Db 595 CGAAACCGCGCTACTACGCTCCGCGACCGCAAAATTAAGACAAATTCACCGATCTTTT 536  
Qy 241 CTCGCGCTACCTCGGCTACCCCGGGAACCGCTTGAGCTGGTGGGAGTGGAGATTTC 300  
Db 535 CTCGCGCTACCTCGGCTACCCCGGGAACCGCTTGAGCTGGTGGGAGTGGAGATTTC 476  
Qy 301 CTGCTCTCCGACAGCCCCCGCCACCTGGGCGAGTGTGAGCTGTGACCTTGC 360  
Db 475 CGTACTCTCGACACCCCGACCCCACTAAACAAATTAATTAAGCTGTGACCTTGC 416  
Qy 361 AGGACGCTGTGAGAGAGAGGCGCTGTGAGACCGCCCGGTGAGAAAGTGGGCTTCA 420  
Db 415 AAAACGCTACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 356  
Qy 421 GCCGCGCTAAAGAGAGAGGCGAGCGTGTGCGCCCGGAGTGTGCGCGCTGTGCGCTT 480  
Db 355 ACCGCGCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 296  
Qy 481 GCTGAGCTCGGCGCTCATGAGGCGACACCGCGCTGTGCGAGGCTCAAGGCGCT 536  
Db 295 ACTAACTCGGCGCTCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 240

RESULT 8  
US-10-363-345A-30994

; Sequence 30994, Application US/10363345A  
; Publication No. US20040234960A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexander Olek  
; APPLICANT: Christian Piepenbrock  
; APPLICANT: Kurt Berlin  
; TITLE OF INVENTION: Method for determining the degree of methylation of defined  
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3  
; FILE REFERENCE: E01/1227  
; CURRENT APPLICATION NUMBER: US/10/363,345A  
; PRIOR FILING DATE: 2003-03-03  
; NUMBER OF SEQ ID NOS: 40712  
; SEQ ID NO 30994  
; LENGTH: 874  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; OTHER INFORMATION: CpG-island No: 30994  
US-10-363-345A-30994

Query Match 36.1%; Score 320; DB 20; Length 874;  
Best Local Similarity 74.8%; Pred. No. 1.4e-85;  
Matches 401; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

Qy 1 CGGCGCAAGAAACAGGAGCGCGCGCCCGCCAGAGAGCGGACATGATTGACCA 60  
Db 100 CGAACCAAAACCAAGAAACCGGACCCCGCAACAAACCAACCAATTAATTAACA 159  
Qy 61 GTTGGGAGGAGCAACCATGCGACCGCTGGGAGGAGCAACGAGTGTGCTGAC 120  
Db 160 ATTACGAAAGAGCAACCATTAACCGGCTAGGAAAGAGCAACGAACTTAATTAAC 219  
Qy 121 CGACTACTGGGATGCTGCGCGCGGAAACCGGACCGCCGACCGCGCATTCACGCC 180  
Db 220 CGACTACTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 279  
Qy 181 CGAGGCGCGTGTGCTGCGCGCGCGCGCGAGTTAGGAGAGTTCACCGTCTTTT 240  
Db 280 CGAAACCGCGTACTACGCTCGGCGCGCGCAAAATTAAGCAAAATTCACCAATCTTTT 339  
Qy 241 CTCGCGCTACCTCGGCTACCCCGGGAACCGCTTGAGCTGGTGGGAGTGGAGATTTC 300

Db 340 CTCGCGCTACCTGACTACACCCGAAACCGCTTGAACTAATAAGCTAATAAGAAATTC 399  
Qy 301 GGTGCTTCGACAGCCCCCGGCCCACTGGGAGAGTGTGACGCTGTGACCTTGC 360  
Db 400 GGTACTTCGACAGCCCCCGGCCCACTAATAAATAAATAAATAAATAAATAAATAAATAA 459  
Qy 361 AGGACGCTGTGAGAGAGGCGCGTGTGACCGCCCGGTGAGAAAGTGGGCTTCA 420  
Db 460 AAAACGCTACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 519  
Qy 421 GCCGCGCTAAAGAGAGGCGAGCGTGTGCGCCCGGAGTGTGCGAGCGCTGTGCGCTT 480  
Db 520 ACCGCGCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 579  
Qy 481 GCTGAGCTCGGCGCTCATGAGGCGACACCGCGCTGTGAGGCTCAAGGCGCT 536  
Db 580 ACTAACTCGGCGCTCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 635

## RESULT 9

US-09-908-975-15549

; Sequence 15549, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: MASSERMAN, Alon  
; APPLICANT: MINTZ, Eli  
; APPLICANT: MINTZ, Liat  
; APPLICANT: FAIGLER, Simchaon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
; FILE OF INVENTION: THAT POPULATE A TRANSCRIPTOME  
; FILE REFERENCE: 36688-0005  
; CURRENT APPLICATION NUMBER: US/09/908,975  
; PRIOR FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/287,724  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: US 60/221,607  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 32337  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15549  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-908-975-15549

Query Match 6.8%; Score 60; DB 10; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8.9e-08;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 565 CCCCTTTCACCTGCTTTTGGAGAAACAGCTGTGAGGCTTTTCTGTATGCTTGT 624  
Db 1 CCCCTTTCACCTGCTTTTGGAGAAACAGCTGTGAGGCTTTTCTGTATGCTTGT 60  
RESULT 10  
US-10-338-110-119  
; Sequence 119, Application US/10338110  
; Publication No. US20040023254A1  
; GENERAL INFORMATION:  
; APPLICANT: Fuhrmann, Jeffrey J.  
; APPLICANT: Romeser, James A.  
; TITLE OF INVENTION: A Method To Assess Quorum Sensing Potential Of Microbial  
; FILE OF INVENTION: Communities  
; FILE REFERENCE: HSR-0056  
; CURRENT APPLICATION NUMBER: US/10/338,110  
; PRIOR FILING DATE: 2003-01-07  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 119  
; LENGTH: 536  
; TYPE: DNA  
; ORGANISM: Artificial Sequence

```
FEATURE:
OTHER INFORMATION: Consensus Sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION: (7)..(7)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (21)..(21)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (33)..(33)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (69)..(69)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (87)..(87)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (213)..(213)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (226)..(226)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (261)..(261)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (273)..(273)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (347)..(347)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (357)..(357)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (398)..(398)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (419)..(419)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (509)..(509)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (517)..(517)
OTHER INFORMATION: n is a, c, g, or t
US-10-338-110-119
```

Query Match 6.3%; Score 56; DB 17; Length 536;  
Best Local Similarity 15.2%; Pred. No. 3.3e-06;  
Matches 61; Conservative 194; Mismatches 144; Indels 1; Gaps 1;

```
Oy 118 GAGCGACCTGAGTCTGCGCCGAGACCCGAGCCGCGCATTCAC 177
    ::|||::|::|::|::|::|::|::|::|::|::|::|::|
Db 129 SSKTCACTTSBTSBGSHVKGWVYVGBDHTSKYGGHTGSCSGCTGHTMCCBAC 188
Oy 178 GCCCGAGGCGCGTGTGCGCTCCGCGCGCGCAGTTACGAGATTCAACGGTCTT 237
```

```
Db 189 SACVBDSBCEWAYCTGCTBBAARKVNSTYTYTRS-BMYNYSBKCBVSRVMMYRCCSM 247
    :|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 238 TTTCGCGCCACTCGGCTACCCCGGGAACCGCTTGAGCTGTGAGTGGCGGA 297
    |||::|::|::|::|::|::|::|::|::|::|::|::|
Db 248 VSVATCCVKMSRTYTGGAHTKWCNGMTWCRSCRCRCRHSAMTKGSSRTGMV 307
Oy 298 TTCGTGCTCTCCGACAGCCCGGCCCACTTGAGGAGAGTGTGACGCTGTACCTT 357
    ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 308 KMRTYTSKSDVMHBCSTSVARWEMGSDSDSYWVBSBGNGCVRKHSVNTSRTSGSVR 367
Oy 358 CGCAGGAGCGCTGTGAGAGAGGCGCGTGTGACCGCCCGTGTGAAGAATGGGCTT 417
    ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 368 SRYSMHCRBSRSXATGVMGGSYKWTWCVDNMKSAMCGBRVTWDSBSHNSANGSVYGG 427
Oy 418 CCAGCCGGGCTAAAGAGAGGAGGAGCGTGCAGCCGCGGAGCTGAGCGCTGTGAGC 477
    ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 428 BYCSYEMHSVMMVHRRBRBDMHRTSSISGSGSTYTGAYCRVADCRMDSCVYVM 487
Oy 478 CTTCGTGAGCTCGCGGCTCATGAGGAGACCGGCGCTGG 517
    ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 488 SVHBYKYWSCYKYRSNMHYCVNRHBBGVMMVKYBVVG 527
```

```
RESULT 11
US-10-437-963-2495
Sequence 2495, Application US/10437963
Publication No. US2004012343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 2495
LENGTH: 989
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_102262C.1
US-10-437-963-2495
```

Query Match 6.2%; Score 55.2; DB 19; Length 989;  
Best Local Similarity 48.1%; Pred. No. 7.3e-06;  
Matches 156; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

```
Oy 68 GAGCGACACCATGATCGGACCGCGCTGCGGAGCGACCGAGCTGTGTCGCCGATAC 127
    |||::|::|::|::|::|::|::|::|::|::|::|::|
Db 279 GACGACACCGTGCCTGCGACCACTCCACACTTCAGACGCGTCATCTCCCTCGCG 338
Oy 128 CTGGGATCTGCGCGCGGGAACCGGCAACCCCGAAGCGCGCATCCAGCCCGAGGCC 187
    |||::|::|::|::|::|::|::|::|::|::|::|::|
Db 339 CTGCTGCTTAAAGCGCGCGCGCGCGCGCTTCGACGTGGCGCGGTCGACGAGCCGC 398
Oy 188 GCCGCTGCGCTCGCGCGCGCGCGAGGTTACGAGATTACCGGTCCTTTCTCCGCC 247
    |||::|::|::|::|::|::|::|::|::|::|::|::|
Db 399 GCGTGTGACGCTGTGCGCGCGCGCGCTGCGCGAGCAGTGTCTTCCACCTC 458
Oy 248 TACCTGCGTACCCCGGAAACCGCTTGAGCTGTGCGCTGTATGGCGGATTCGTGCTC 307
    |||::|::|::|::|::|::|::|::|::|::|::|::|
Db 459 CACTCGGAGACCAAGCGGCGGTGAAGGAGATACCACTGTGCTGACAGGCGTGTG 518
Oy 308 TCCGACAGCCCGGCGCCCACTGGGAGAGTGTGACGCTGTGACCTTGGCAGGAGCG 367
    |||::|::|::|::|::|::|::|::|::|::|::|::|
```



Query Match 6.2%; Score 54.8; DB 20; Length 1743;  
 Best Local Similarity 48.1%; Pred. No. 1.2e-05;  
 Matches 226; Conservative 0; Mismatches 232; Indels 12; Gaps 2;

QY 73 CACCAACATGCGGACCCCGCTGCGGAGCGACCCAGCTGTGCTGCGCCGACTTACTGGG 132  
 DB CGCCCGCTGCGGACCCCGCGCCGACGACCGCGCTGCGGACCCGCTGAGCTTACTGGG 898  
 QY 133 GTACTGCGCGGAGAACCCCGGACCCCGGAGCGGCGCATCCAGCCCGAGGCGCGCT 192  
 DB GTCCCGCGCGGAGCGGCG 958  
 QY 193 GCTGCGCTCCGCT 252  
 DB GCGCGACCG 1018  
 QY 253 CCGCTACCCCGGAAACCGCTTGCAGCTGCGCGCTGATGCGGATTCCTGCTCTCCGA 312  
 DB 1019 GCGCATGACGCT-----CGTGGACGGCGAGCGCTGCGGAGCTCGCGAGTGTCTCG 1072  
 QY 313 CAGCCCGCT 372  
 DB 1073 CGAGCTGAGAGAGGCT 1132  
 QY 373 GAGAGAGGCGCGCGCTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTAA 432  
 DB 1133 GAAGCGCGTGGAGCAGCAG-----CTGAGATGAGCGCGCGCGCGCGCGCGAGT 1186  
 QY 433 GAGAGAGAGGCT 492  
 DB 1187 GTGGAGGAGGTTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 1246  
 QY 493 GCTCATGCGGCGAGCAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGATG 542  
 DB 1247 GCGCGTGGAGCGCGGACACCGGGGAGAGGTTGCGGAGCTCGCGCGCGAGC 1296

RESULT 15  
 US-10-156-761-7129  
 ; Sequence 7129, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 7129  
 ; LENGTH: 954  
 ; TYPE: DNA  
 ; ORGANISM: Streptomyces avermitilis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(954)  
 US-10-156-761-7129

Query Match 6.0%; Score 53.4; DB 15; Length 954;  
 Best Local Similarity 47.0%; Pred. No. 2.5e-05;  
 Matches 165; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 23 GCCCGCCCCCGACGAGGCGCGGACCATGTTGACCAAGTTCGAGGCGCGACCAACATG 82  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 403 GCCGAAATCTTGTGATACCGGACCGGCGCTGAGCGAGGAGCCGACCGCTCCGCCACG 462  
 QY 83 GCCGACCGCTGCGGAGCGACCGAGCTGTGCTGCGGACTACCTGAGGTTACTGCGCC 142  
 DB 463 CTCGCGCGCTTCAAGGCGCGGACCAAGGAGGAGCACACAGGTTCTGACCTGACTGGGCG 522  
 QY 143 CGGAAACCGGACCGCGCGGACCGCGCATCCAGCGCGGAGCGCGCTGCGCTCC 202  
 DB 523 CGGATGTTCTGGCAGGACCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 582  
 QY 203 CGGCGCGCAGGTTACGCAATTCACCGGTCCTTTTCTCGGCTTACTGCGCTACCC 262  
 DB 583 GCGACCTGCGCGCTGCGCACTGACGAGCGAGATGCGCACCGGCGTACGCGAACCC 642  
 QY 263 GGGAAACCGCTTCAAGCTGTGAGCGCTGAGCGGATTCGCTCTCCGACAGCGCGCGC 322  
 DB 643 CGGCGCTGCGCGGAGCGCTTATGAGGCGGCGCGGAGCTGCGCTGTGTTAAACAGGCG 702  
 QY 323 CCCACCTGAGGCGAGAGTGAGCGCTGTGACCTTCCGAGGAGCGCTGCTG 373  
 DB 703 CCGAAGGAGTGTCTGCGCGGTCACCGGACCGGACACCGCGGAGGTCGCG 753

Search completed: June 7, 2005, 18:04:16  
 Job time : 839.797 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 06:20:59 ; Search time 5392.65 Seconds  
(without alignments)  
6260.935 Million cell updates/sec

Title: US-10-071-174A-1

Perfect score: 887

Sequence: 1 cgggcacgaagaaccacgcga.....ctctccttgagtgaaagaa 887

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :	EST :
1: gb_est1.*	
2: gb_est2.*	
3: gb_est3.*	
4: gb_est4.*	
5: gb_est5.*	
6: gb_est6.*	
7: gb_est7.*	
8: gb_est8.*	
9: gb_est9.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	289.4	32.6	715	6	CA777633 ip19a09.Y
2	236.2	26.6	626	4	BG080862 H3058C03.Y
3	235.4	26.5	875	4	CO798715 AGENCOURT
4	235.2	26.5	649	7	CF915355 B0978E08-
5	233.4	26.3	772	7	CO814648 AGENCOURT
6	231.6	26.1	736	6	BY735639 BY735639
7	224.2	25.3	588	6	CA557550 K0229B10-
8	220.8	24.9	758	7	CO799635 AGENCOURT
9	213.6	24.1	821	7	CO806531 AGENCOURT
10	206.4	23.3	524	6	CA560564 K0274C06-
11	191.6	21.6	763	4	CO808612 AGENCOURT
12	180.2	20.3	706	1	AA005293 zh93a11.r
13	169.8	19.1	478	6	CA561864 K0293B04-
14	169.8	19.1	478	6	CA561864 K0293B04-
15	167	18.8	535	5	BX119239 BX119239
16	157.6	17.8	448	6	CA560084 K0267E03-
17	153	17.2	545	4	BG071824 H3103E12-
18	138.2	15.6	721	6	BY756023 BY756023
19	137.6	15.5	860	7	CO811531 AGENCOURT
20	136.8	15.4	486	1	AT505208 vq73g11.x
21	136.8	15.4	493	7	CO797157 AGENCOURT
22	131.8	14.9	493	7	AT181346 wj33g03.x
23	123.4	13.9	387	5	BY029373 BY029373
24	119.8	13.5	422	6	CA561120 K0282F09-

C 25	117.2	13.2	638	4	BM228756 K0267E03-
26	115	13.0	380	5	BY029358 BY029358
27	114.8	12.9	130	1	AA098865 zka4f02.b
28	107.6	12.1	440	5	BY029830 BY029830
29	107.2	12.1	352	1	AA426934 vE22b11.b
30	106.6	12.0	399	1	AA623872 vq73g11.b
31	99	11.2	363	4	BG084666 H3103E12-
32	97.4	11.0	615	4	BM230283 BM230283
33	96.8	10.9	339	1	AT604980 vE22b11.x
34	84.4	9.5	313	1	AT467076 vD09e08.x
35	69.2	7.8	925	9	AL053013 Drosophila
36	63.8	7.2	925	9	CNS0091P CNS0091P
37	62.8	7.1	330	6	CA556079 CA556079
38	60.8	6.9	661	9	CE629081 tigr-998-
39	60.4	6.8	1267	9	CA765590 SAIL_258
40	60.4	6.8	1552	9	AG430101 Mus muscu
41	60.2	6.8	606	5	BP111813 BP111813
42	59.6	6.7	935	9	AL066051 Drosophila
43	59.4	6.7	1070	8	BZ557810 paca1-60
44	58.4	6.6	931	7	CK412737 AUF Ip64I
45	58.2	6.6	706	1	AT455155 AT455155

## ALIGNMENTS

RESULT 1  
CA777633/c 715 bp mRNA linear EST 03-DEC-2002  
LOCUS ip19a09.Y1 HR85 1slet Homo sapiens cDNA clone IMAGE:6217625 5'  
DEFINITION similar to TR:Q920F3 Q920F3 BCL-2 HOMOLOG. ; mRNA sequence.

ACCESSION CA777633.1 GI:26015508  
VERSION CA777633  
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Melton, D., Brown, J., Keny, G., Permutt, A., Lee, C., Keestner, K.,

Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Page, D., Wylie, T., Martin, J., Blisat, A.,

Schmitt, A., Theising, B., Ritzer, B., Ronko, I., Bennett, J.,

Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvilli, R.,

Williams, T., Jackson, Y., and Bowers, Y.

Unpublished (2000)

Endocrine Pancreas Consortium

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@molb.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:

Washington University Genome Sequencing Center for information on

obtaining a clone please contact: Dr. Hiroshi Inoue

(hinoe@im.wustl.edu)

Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gibco

High quality sequence stop: 392.

Location/Qualifiers

1..715

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6217625"

/rname\_type="purified pancreatic islet"

/lab\_host="DH10B"

/clone\_id="HR85 1slet"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:

Note: Site\_2: XhoI; cDNA made by oligo-dT priming.

Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permut Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

Query Match 32.6%; Score 289.4; DB 6; Length 715;  
Best Local Similarity 93.5%; Pred. No. 1.8e-60;  
Matches 302; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY CCCCTTTCACCTGCTTTTGGAGAAACAGCTGCTTCTTCTGATGCTTGT 624  
DB CCCCTTTCACCTGCTTTTGGAGAAACAGCTGCTTCTTCTGATGCTTGT 648  
QY 625 AACACAGCTTCACTTATCTCTGACACGATTTATGATTTTAAACCTTTAACCCG 684  
DB 647 ACCACAGCTTCACTTATCTCTGACACGATTTATGATTTTAACTTTTACCCG 588  
QY 685 CTTCACCTGCTTCACTGACCACTAAATGACATGCTGACAAACAGACTGAGG 744  
DB 587 CTTCACCTGCTTCACTGACCACTAAATGACATGCTGACAAACAGACTGAGG 528  
QY 745 AAAGACCTTCCCTCCACCCAGACGCTTTTATCTGAATGACATCAAGAGTCTGAGGTG 804  
DB 527 AAAGACCTTCCCTCCACCCAGACGCTTTTATCTGAATGACATCAAGAGTCTGAGGTG 468  
QY 805 GTGATTTGGCCAGTGTTTTAACTTGTGACACAGTCTGAGTGTGAGACAGAAATGCAA 864  
DB 467 GTGATTTGGCCAGTGTTTTAACTTGTGACACAGTCTGAGTGTGAGACAGAAATGCAA 408  
QY 865 TGGCTCTCTCTGAGTGAAGA 887  
DB 407 TGGCTCTCTCTGAGTGAAGA 385

RESULT 2  
LOCUS BG080862 626 bp mRNA linear EST 17-DEC-2003  
DEFINITION H3058C03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
H3058C03 5', mRNA sequence.  
ACCESSION BG080862  
VERSION BG080862.2 GI:40014930  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 626)  
Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grahovac, M.J., Pantano, S., Sano, Y., Fiao, Y., Nagataj, K., Doi, H., Wood, W.H., III, Becker, K.G., and Ko, M.S.H.  
Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray  
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)  
20381348  
10922068  
On Jan 26, 2001 this sequence version replaced gi:12563419.  
Other ESTs: H3058C03-3  
Contact: George J. Kargul  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
This set has been freely distributed to the community. Please visit <http://igsun.grc.nia.nih.gov/cDNA/15k.html> for details.  
Plate: H3058 row: C column: 03  
Seq primer: -21M3 Reverse  
High quality sequence stop: 626  
POLYA=No.

FEATURES

Location/Qualifiers

source

1..626  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:H3058C03-5"  
/db\_xref="taxon:10090"  
/clone="H3058C03"  
/sex="Clones arrayed from a variety of cDNA libraries"  
/dev\_stage="Clones arrayed from a variety of cDNA libraries"  
/lab\_host="DH10B"  
/clone\_id="NIA Mouse 15K cDNA Clone Set"  
/note="Vector: pSPORI1; Site 1: SalI; Site 2: NotI. This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos), and E12.5 female mesonephros/gonad, and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match 26.6%; Score 236.2; DB 4; Length 626;  
Best Local Similarity 65.1%; Pred. No. 2.4e-47;  
Matches 407; Conservative 0; Mismatches 193; Indels 25; Gaps 3;

QY 86 GACCCGCTGCGGAGCGCAGACCTGCTGCGCCGACTACCTGCGGCTGCGG 145  
DB 18 GACCCGCTGCGGAGCGCAGACCTGCTGCGCCGACTACCTGCGGCTGCGG 77  
QY 146 GAACCCGCTGCGGAGCGCAGACCTGCTGCGCCGACTACCTGCGGCTGCGG 205  
DB 78 GAGCGGAGCACCAGCGCCGACCGCCGCTGCTGCGGAGCGGCTGCTGCTG 137  
QY 206 GCGCGAGCTTACCGGAGATTCACCGGCTTTTCTCGGCTACCTGCTACCGG 265  
DB 138 ACTAGGAGATTCACCGGAGATTCACCGGCTTTTCTCGGCTACCTGCTACCGG 197  
QY 266 AACGCTTGCAGCTGCTGCGGCTGATGCGGATTCGCTCTCGGACAGCCCGCC 325  
DB 198 AATGCGCTGAGCTGCTGTAACAGATGCAATTAAGTGTCTCTCAAGACCAAGACTTC 257  
QY 326 ACTGCGGAGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 385  
DB 258 AGCTGAGCCCACTGCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 317  
QY 386 CTGCTGAGCGGCGGCTGGAAGATGCGGCTTCCAGCGCGGCTTAAGAGACAGAGG 445  
DB 318 TACATGCTGCTGCAACAGACAGACAGACAGACAGACAGACAGACAGACAG 359  
QY 446 GACCTGCGCCGCGGAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502  
DB 360 ATAAGTACCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419  
QY 503 CAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562  
DB 420 CGGACCTGCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479  
QY 563 ACCGCTTTCAGCTGCTTGTGAGAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622  
DB 480 AATCTTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539

QY 623 TTAAACAGACGCTTATCTGACACGATTATATGATTTAAACTTTAAC 682  
DB 540 TTGGACAGCGCATCTTTTATCTGGAAACGTTAT-----AAGTTAAATTTTAAAG 595  
QY 683 CGCTTCTACCTGCGCACTGTGAC 707  
DB 596 CGATCTTACTTACTTACTGTAAC 620

RESULT 3  
LOCUS CO798715 875 bp mRNA linear EST 05-AUG-2004  
DEFINITION AGNCOURT 30259012 NIH\_MGC 256 Mus musculus cDNA clone  
IMAGE:30935405 5', mRNA sequence.  
ACCESSION CO798715  
VERSION CO798715.1 GI:50986895  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 875)  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cga@bbs-rcmail.nih.gov  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDAM1189 row: h column: 06  
High quality sequence stop: 659.  
Location/Qualifiers  
1. 875  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30935405"  
/lab\_host="DH10B TONa"  
/clone\_1lb="NIH\_MGC 256"  
/note="Organ: oocyte; Vector: pExpress-1; Site: 1: EcoRV,  
Site 2: NotI; cDNA was primed using oligo-dT primer:  
5'-pGACTAGTTCTAGATCGGAGCGCGGCC(7)25-3' and cloned into  
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb  
resulted in an average insert size of 1.2 kb. This is a  
primary library (normalized primary library is NIH\_MGC 257)  
and was constructed by Express Genomics (Frederick, MD).  
Note: this is a NIH\_MGC library"

ORIGIN  
Query Match 26.5%; Score 235.4; DB 7; Length 875;  
Best Local Similarity 64.2%; Pred. No. 4,1e-47;  
Matches 414; Conservative 0; Mismatches 206; Indels 25; Gaps 3;

QY 86 GACCCGCGCGGAGACGACGAGCTGTTCGCGCGCACTACTGGGGTACGCGCCGG 145  
DB 27 GACCCAGTCAGTAAGACGACGAGCGGCTGTCTGACTATATTTCTTTCGCGCAGG 86  
QY 146 GAACCCGCGACCCCGGACCGGCGCCATTCACGCGCGCGCGCGTGTGCGTCCGG 205  
DB 87 GAGCCGACACCCCGGACGACCGCGCCCACTGTCTGAGCGCGCGCTTGTGCTCTGG 146  
QY 206 GCGCCAGGTTACGCGAGATTTCACCGGTCCTTTTCTCGGCTTACTCGGCTAACCCCGG 265  
DB 147 ACTAGGAGATTCACAGAGACCAAGAAATTTTCTCTCTCTGCGAAGCGCGGCG 206

QY 266 AACCGCTTCAGCTGTGCGCGCTGATGCGGATTCCTGCTCTCCGACAGCCCGGCC 325  
DB 207 AATGCCCTGAGCTGTGTGTAACAGATGGCAATAGTTGCTCTCCAAAGACCAAGACTTC 266  
QY 326 ACCTTGGGACAGTGTGTAACGCTCTGACCTTTCCGAGGAGCGCTGTGAGAGAGCGG 385  
DB 267 AGCTGAGACCAACTGTGTATGCTCTGCGCTTCGCGGGAGCGCTTATATCAAGGCCCT 326  
QY 386 CTGCTGACCGCCCGGTGAGAGATGAGGGCTTCCAGCGCGGCTTAAAGAGACGAGGC 445  
DB 327 TACATGCTGTCAACAGAAAG-----GGAGTCTGGGATGCTGTC 368  
QY 446 GACGTGCCCGGAGACTGCGCAGCGCTGTGCGCTTGTGAGCTCGCGCTCAATGCGG--- 502  
DB 369 ATAGTACCCGACAGACTGCTGTCTCATATGTAACCTTCTGTATATCTGTCTCATAGGGGGT 428  
QY 503 CAGACCGCGCGCTGCTGACAGGCTCAAGGCGGCTGTGAGATGCGCTTTGTCACTTTCAG 562  
DB 429 CGGACCGCGCGCAGGCTGAGAGGCTCTCGCGCGGCTGTGAGATGCGCTTTGCGCTTCAAG 488  
QY 563 ACCCGCTTTCAGCTGCGCTTTTGGAGAAACAGCTGCTCCAGGCTTTCGTGATGCTG 622  
DB 489 AATCTTTCAGCTGCGCTTCTGAGAAAGATTGCTGATTCAGGCTTTCGTGAGGCTTC 548  
QY 623 TTAAACAGACGCTTATCTCTGACACGATTATATGATTTAAACTTTAAC 682  
DB 549 TTGGACAGCGCATCTTTTATCTGGAAACGTTAT-----AAGTTAAATTTTAAAG 604  
QY 683 CGCTTCTACCTGCGCACTGTGACCACTAAATGACAGATGTG 727  
DB 605 CGATCTTACTTACTTACTTACTGTAACCTCTAAGAACTTCGCG 649

RESULT 4  
LOCUS CF915355 649 bp mRNA linear EST 05-NOV-2003  
DEFINITION B0978E08-5 NIA Mouse Unfertilized Egg cDNA Library (long 1) Mus  
musculus cDNA clone NIA:B0978E08 IMAGE:30479767 5', mRNA sequence.  
ACCESSION CF915355  
VERSION CF915355.1 GI:38186557  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 649)  
Piao, Y., Ko, N.T., Ijm, M.K. and Ko, M.S.H.  
Construction of long-transcript enriched cDNA libraries from  
submicrogram amounts of total RNAs by a universal PCR amplification  
method  
Genome Res. 11 (9), 1553-1558 (2001)  
21429098  
MEDLINE  
PUBMED  
COMMENT  
JOURNAL  
LABORATORY  
CONTACT  
DAWOOD B. DUDKULA  
LABORATORY OF GENETICS  
NATIONAL INSTITUTE ON AGING/NATIONAL INSTITUTES OF HEALTH  
333 CABELL DRIVE, SUITE 4000, BALTIMORE, MD 21224-6820, USA  
EMAIL: cdna@gsun.grc.nia.nih.gov  
PLATE: B0978 ROW: B COLUMN: 08  
SEQ PRIMER: M13 Reverse  
High quality sequence stop: 649  
POLY(A)-No.  
Location/Qualifiers  
1. 649  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/db\_xref="taxon:10090"  
/clone="NIA:B0978E08 IMAGE:30479767"  
/dev\_stage="Unfertilized Egg"  
/lab\_host="DH10B"  
/clone\_1lb="NIA Mouse Unfertilized Egg cDNA Library (long

1)"  
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;  
Site 2: NotI; Mouse cDNA project by the Laboratory of  
Genetics, National Institute on Aging (NIA), Intramural  
Research Program, NIH (<http://nisan.gtc.nia.nih.gov/cDNA>).  
This is a long-transcript enriched cDNA library (Ref.  
Genome Res. 11: 1553-1558 (2001). [PMD: 11544191]). Total  
RNAs were extracted from a pool of 148 unfertilized eggs.  
Double-stranded cDNAs were synthesized with an Oligo(dT)  
primer (Invitrogen):  
5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3',  
treated with T4 DNA polymerase, and purified by  
ethanol-precipitation. The cDNAs were ligated to  
lone-linker l1-SalI, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Taq polymerase (Takara) with a primer SalI-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid  
vector. The DH10B E. coli host was transformed with the  
ligation mixture by the standard chemical method. The  
average insert size is about 2.5 kb. The library was  
constructed by Yulan Piao."

## ORIGIN

Query Match 26.5%; Score 235.2; DB 7; Length 649;  
Best Local Similarity 64.4%; Pred. No. 4,3e-47;  
Matches 412; Conservative 0; Mismatches 203; Indels 25; Gaps 3;

86 GACCCGCTGCGGAGACGACCGAGCTGTGTGCGCGACTACCTGGGGTACTGCGCCGG 145  
18 GACCCACTGCAATGAAGACGACTAGCGGCTGTGTGACTATATTTCTTCCGACGCG 77  
146 GAACCCGCGACCCCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 205  
78 GAGCGGACACCCCGAGGCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 137  
206 GCGCGCAGGTTACGCGAGATTCACCGGCTCTTTTCTCGGCTACTCGGCGTACCCCGGG 265  
138 ACTAGGCGAGATCAAGAGAGACCAAGAAATTTTCTCTCTTCTGCGAAACCGGCGG 197  
266 AACCGCTTGAAGTGTGCGCTGATGCGGATTCGCTGCTCCGACAGACGCCGCGCCCG 325  
198 AATGCCCTGAGCTGTGTAACAGATGAGATAGTGTCTTCCAAAGACCAAGACTTC 257  
326 ACTTGGGCGAGAGTGTGACGCTGCTGACCTTTCGAGGAGCGCTGCTGAGAGAGGCGCG 385  
258 AGCTGAGGCAACTGTGTGCTCTGCGGCTTTCGCGGAGCGCTTATGATCAAGGCGCT 317  
386 CTGCTGACCGCGCGGTGGAAGATGGGCGCTTCCAGCGCGGCTAAAGAGAGAGGAGCG 445  
318 TACATGCGCTGTCAAGCAAGAA-----GGGATCTGGGGAATGCTGTTC 359  
446 GACCTGCGCGGAGACTGCCAGCGCGCTGTGAGCTTGTGAGCTCGCGGCTCAATGAGG-- 502  
360 AATAGTGAACCGAGACTGCTGTCTCATATGAACTTTCTGTATATCTGTCTATGGGCGCT 419  
503 CAGCACCGGCGCTGCTGAGGCTCAGGGCGGCTGGAGTGGCTTTTGTCACTTCTTCAGG 562  
420 CGGCAACCGGCGAGGCTGAGGCTCTCGGCGGCTGGAGTGGCTTTTGTGCGCTTCTTCAG 479  
563 ACCCGCTTTCAGCTGGCTTTTGGAGAAACAGCTGGTCCAGGCTTTTGTCAATGCTTG 622  
480 AATCTTTTACCGCTCGGCTTCTGAGAAATGCTGATTCAGGCTTTTGTCTAGGCTTC 539  
623 TTAAACAAGCCTTCAATTTATCTGTGACACGATTAATTAATGATTTTAAACCTTTTAC 682  
540 TTTGGAAGAGCATCTTTTATCTGGAACGTTAT-----AAGTTAAATTTTAAAG 595  
683 CGCTTCTACCTGCGCAACTGTGACCAACTAAATGACAGT 722  
596 CGATCTTACTACTTACTGTGAACCTCTTAAGAGCAAT 635

RESULT 5  
LOCUS COB14648 772 bp mRNA linear EST 06-AUG-2004  
DEFINITION AGNCOURT 30246837 NIH\_MGC 256 Mus musculus cDNA clone  
IMAG:30937746 5', mRNA sequence.  
ACCESSION COB14648  
VERSION COB14648.1 GI:51033274  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.  
1 (bases 1 to 772)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics / NIH  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
CDNA Library Preparation: Express Genomics  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUMI)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LUMI at:  
<http://image.llnl.gov>  
Plate: NDM1195 row: 1 column: 19  
High quality sequence start: 5  
High quality sequence stop: 666.  
location/Qualifiers

## FEATURES

source

1..772  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAG:30937746"  
/lab\_host="DH10B Torus"  
/clone\_lib="NIH\_MGC\_256"  
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV;  
Site 2: NotI; cDNA was primed using oligo-dT primer:  
5'-pGACTAGTCTTCAATGATCGGAGCGGCGGCGGCTT-3' and cloned into  
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb  
resulted in an average insert size of 1.2 kb. This is a  
primary/library (normalized primary library is NIH\_MGC 257)  
and was constructed by Express Genomics (Frederick, MD).  
Note: this is a NIH\_MGC library"

## ORIGIN

Query Match 26.3%; Score 233.4; DB 7; Length 772;  
Best Local Similarity 63.9%; Pred. No. 1.2e-46;  
Matches 412; Conservative 0; Mismatches 208; Indels 25; Gaps 3;

86 GACCCGCTGCGGAGACGACCGAGCTGTGTGCGCGACTACCTGGGGTACTGCGCCCG 145  
42 GACCCACTGCAATGAAGACGACTAGCGGCTGTGTGACTATATTTCTTCCGACGCG 101  
146 GAACCCGCGACCCCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 205  
102 GAGCGGACACCCCGAGACCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 161  
206 GCGCGCAGGTTACGCGAGATTCACCGGCTTTTCTCGGCTACTCGGCGTACCCCGGG 265  
162 ACTAGGCGAGATCAAGAGAGACCAAGAAATTTTCTCTTCTTGTGGAAGCGGCGGCG 221  
266 AACCGCTTGAAGTGTGCGCTGATGCGGATTCGCTGCTTCCGACAGCCCGGCGCCG 325  
222 AATGCCCTGAGCTGTGTAACAGATGAGATAGTGTCTTCCAAAGACCAAGACTTC 281  
326 ACTTGGGCGAGAGTGTGACGCTGCTGACCTTTCGAGGAGAGCTGCTGAGAGAGGCGCG 385



Db 282 AGCTGAGCAACGCTGATGCTCTGCGCTTCGCGGAGACGCTTAATGAAACAGGCCT 341

Qy 386 CTGAGACCGCCCGGTGAAGAAGTGGGCTTCACCGCGCTAAAGACGAGAGGC 445

Db 342 TACATGCGCTGTACACAGAGA-----GGGATCTGGGGATCGTGC 383

Qy 446 GACGTCGCGCGGAGCTGACGCGCTGAGCTTCGCTGAGCTCGCGCTCAATGGG--- 502

Db 384 ATAGAGACCGAGATGCTGTCTCATAGGAACCTTCGTATATATCTGTCTATGGGGCT 443

Qy 503 CAGCAGCGCGCTGCTGAGGCTGACGCGCGCTGGATGCGCTTTGTCATCTTTCAG 562

Db 444 CGGACCGCGCGAGCTGAGAGGCTCGCGCGCTGGATGCGCTTTGCGCTTTCAG 503

Qy 563 ACCCGCTTTCACGCTGCTTTTGGAGAAAGCTGCTGCTTCAGGCTTTCTGCAAGCTTG 622

Db 504 AATCTTTACGCTGCTGCTTGGAGAAAGATGCTGATTCAGGCTTTCTGCAAGCTTG 563

Qy 623 TTAACACAGCGCTTATTTATCTGACACGATTTATGATTTTAAACCTTTAAC 682

Db 564 TTTCACACAGCGATCTTTTATCTGGAACCTTTAT-----AAGTTTAAATTTTAAAG 619

Qy 683 CGCTTCTACCTGCGCACTGTGACCACTAAATGAACAATGTGTG 727

Db 620 CGATCTTACCTACTACCTGTGACCTCTTAAAGAACTTCTGGG 664

RESULT 6  
BY735639 736 bp mRNA linear EST 17-DEC-2002

LOCUS BY735639 RIKEN full-length enriched, 1 cell embryo Mus musculus

DEFINITION CDNA clone I0C0031F16 5', mRNA sequence.

ACCESSION BY735639

VERSION BY735639.1 GI:271148766

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 736)

AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bredt, D., Brusic, V., Chochua, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Mik, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pereira, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Vetraro, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyszynski, B., Yanagisawa, M., Yang, I., Wells, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carinci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, O., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 22354683

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Saito-cho, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carinci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, O., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse libraries.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

source Location/Qualifiers

1..736

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="I0C0031F16"

/cell\_type="1 cell"

/dev\_stage="1 cell embryo"

/clone\_lib="RIKEN full-length enriched, 1 cell embryo"

## ORIGIN

Query Match 26.1%; Score 231.6; DB 6; Length 736;  
Best Local Similarity 63.2%; Pred. No. 3.4e-46;  
Matches 463; Conservative 0; Mismatches 242; Indels 28; Gaps 6;

Qy 86 GACCGCTGCGGAGGACGAGCTGTGCGCGCACTACCTGCGCTACCTGCGCGCG 145

Db 29 GACCCATGCAATGACGACATGACGCTGCTGTGCTCAATATTTCTTGCGCGAG 88

Qy 146 GAACCGGACACCGCGAGCGCGCATTCACGCGCGAGCGCGCTGCTGCTGCGCG 205

Db 89 GAGCGGACACCGCGAGCGCGCATTCACGCGCGAGCGCGCTGCTGCTGCTGCT 148

Qy 206 GCGCGGAGTTACGCGAGATTACCGGCTCTTTTCTCGCGCTTACCTGCGCTAC 265

Db 149 ACTAGGACGATTCAGACGAGACCAAGAAATTTTCTCTCTCTGCGAAGCGCGG 208

Qy 266 AACCGCTTGCAGCTGTGCGCGCTGATGCGCGATTCGCGCTCTCGAAGCGCGG 325

Db 209 AATGCGCTGAGCTGTGTAACAAGATGCAATATGTTCTCTCCAAAGACCAAGATTC 268

Qy 326 ACTTGGGCGAGAGTGTGACCTGCTGCTTCCAGGAGAGCTGCTGAGAGAGGCGG 385

Db 269 AGCTGAGCCAACTGTGATCTCTGCGCTTCCGCGGAGAGCTTATGAACAAGGCGCT 328

Qy 386 CTGCTGACCGCGCGGTGAGAGAGTGGGCTTCCAGCGCGCTTAAAGAGAGAGGCG 445

Db 329 TACATGCGCTGTCAAGACAGAGA-----GGGATCTGGGAATCGTGTG 370

Oy	446	GACGCTGCGCCGGGACCTGCGACGCGCTGGAGCCTTGCTGAGCTCGCGGCTCATAGGGG---	502
Db	371	ATATGATACCCCGAGACTGCTGTCTCAATAGTAACTTTCTGTAAATCTGCTCAATGGGGCCT	430
Oy	503	CAGCACCGCGCTGCGCTGAGGCTCAGGGCGGCTGGATGCTTTGTCTTCACTTCTTCAAG	562
Db	431	CGGCACCGGCGCAGGCTGGAAGGCTCTGCGCGGCTGGAAATGGCTTTTGGCGCTTCTTCAAG	490
Oy	563	ACCCCTTTTCCACTGGCTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCAATGCTTG	622
Db	491	AATCTTTTAACTGCTCGGCTTTCGGAGAAAGATTGCTGAATTCAGGCTTTTCTGTCAAGCTTTC	550
Oy	623	TTAAACAAGCGCTCATTTATCTCGAGACAGATATATGATGTTTTTAAACTTTTAAACC	682
Db	551	TTTTCG-AACGACATCTTTTATTATCTGGAACCTTTAT---AAGTTTAAATTTTAAAG	605
Oy	683	CGCTTTTCACTGGCCCACTGTGACCAACTTAATGACAGATGTGTAGAACAAAGACTGAG	742
Db	606	CGATTTTACCTACTCACTGTGAAACCTCTAAGACAAATTCG-GGGTATTGGGATAGAA	664
Oy	743	GGAAAGCACTTCCCCCAACCCCAAGCTTTTATCTGAATGATATCAAGAGTCTGTAGG	802
Db	665	CTGAAGAAGTCTCTCAGCCTAGACAACTTTTACTGTGATGC-TAGGTGAGAGTCTGGGT	723
Oy	803	TGCTGATTTTGGCC	815
Db	724	TNGTATGGGGCC	736

RESULT	7
CAS57550	
LOCUS	588 bp mRNA linear EST 19-NOV-2001
DEFINITION	K0229B10-5N NIA Mouse Unfertilized Egg cDNA Library (long) Mus musculus cDNA clone NIA.K0229B10 IMAGE:30047445 5', mRNA sequence.
ACCESSION	CAS57550
VERSION	CAS57550.1 GI:25102015
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 588)
AUTHORS	Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Steagy,C.A., Martin,P., Attha,K., Tanaka,T. and Ko,M.S.H.
TITLE	Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library (long)
JOURNAL	Unpublished (2001)
COMMENT	Other ESTs: K0229B10-3 Contact: Dawood B. Dudekula Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@igsun.grc.nia.nih.gov Plate: K0229 row: B column: 10 Seq primer: M13 Reverse High quality sequence scop: 588 POLYA=No.

## FEATURES

```

location/Qualifiers
1. 588
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="nlastr:K0229B10-5N"
    /db_xref="taxon:10090"
    /clone="NIA:K0229B10 IMAGE:30047445"
    /tissue_type="Unfertilized Egg"
    /lab_host="DH10B"
    /clone_lib="NIA Mouse Unfertilized Egg cDNA Library
(Long)"
/note="vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2
Not: Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgun.gic.nia.nih.gov/cDNA). This is

```

a long-transcript enriched cDNA library (Ref. Genome Res  
 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were  
 extracted from a pool of 1488 unfertilized eggs.  
 Double-stranded cDNAs were synthesized with an Oligo (dT)  
 primer [inverted:  
 5'-GCACTAGTCTTAGATCGGAGCGGCGCCCTTTTCTTTTCTTTT-3'],  
 treated with T4 DNA polymerase, and purified by  
 ethanol-precipitation. The cDNAs were ligated to  
 Lona-linker IL-Sal4, purified by phenol/chloroform, and  
 separated from free linkers by Centricon 100. Then, the  
 cDNAs were amplified by long-range high fidelity PCR using  
 Ex Taq polymerase (Takara) with a primer Sal4-S. The  
 products were purified by phenol/chloroform and Centricon  
 100. The cDNAs were digested with SalI and NotI enzymes  
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
 The DH10B E. coli host was transformed with the ligation  
 mixture by the standard chemical method. The average  
 insert size is about 2.5 kb. The library was constructed  
 by Yulan Piao (NTA)."

A long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). (PMID: 11544191)). Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo (dt) primer [nucleotides: 5'-pGACATGCTTCAGATCGGACGCGCCCTTTTCTTTTCTTTT-3'], treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-Linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using the Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with *Sal*I and *Nci*I enzymes and cloned into *Sal*I/*Nci*I site of pSPORT1 plasmid vector. The DH10B *E. coli* host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NTA)."

**ORIGIN**

Query Match	25.3%	Score 224.2;	DB 6;	Length 588;
Best Local Similarity	64.8%;	Pred. No. 2.2e-44;		
Matches 375;	Conservative 0;	Mismatches 183;	Indels 21;	Gaps 21;

OY	86	GACCGCTGCGGGGAGCGCAGCGAGCTGTGGTGGCGCACTACCTGGGGTACACGCGCCGG	145
Db	16	GACCACTGCACTGAAACGCATCTAGACGGGTGCTGTCTGACTACATATTTCTTCTGGCAAG	75
OY	146	GAACCCGGCACCCCGAGCCGGCGCCATCCACGCCCGAGCCCGCCGTGCTGCTCGCG	205
Db	76	GAGCGCGAGCACCCCGAGGCCACCGCGCCACGTCTGTGCGAGCGGGCTTCTCTGCTGTGG	135
OY	206	GCCCGCAGGTTACCGCAGATTCAACCGGCTCTTTTCTCCGGCTACCTGGGCTACCCCGGG	265
Db	136	ACTAGGCGAGATCCACGACGAGCGACCCAGAAATTTTTTCTCTCTCTTCGAAAGCCGGGC	195
OY	266	AACGGCTTCGAGCTGTGGCGCTGATGTCGGAATTCGGTCTCCGACAGCGCCCGGCC	325
Db	196	AATGGCTCTGGAGCTGTGTAAACAGATGCAATAGTTGGTCTCTCCAAAGACCAACATCTC	255
OY	326	ACCTGGGGCAGATGTGTGACCGCTGTGTACCTTCGCGAGGACGCTGTGTAAGAGAGGGCTG	385
Db	256	AGCTGGAGCAACTGTGTATGCTCTCGGCTTCGCGGGAGCGCTTATGATCAAGGCCCT	315
OY	386	CTGGTGACCGCCCGGTGGAAAGTGTGGGGCTTCCAGCCGCGGCTAAAGAGCAGAGGGC	445
Db	316	TACATGGCTGTCAAGCAGAGAA-----GGGATCTGGGGAATGTGTG	357
OY	446	GACGTGCGCCGGGAGCTGCACAGCGCCGTGTGGCTTGTAGAGCTCGCGGCTATGGGG---	502
Db	358	ATAGTGAATCCAGAGCTGTGTCTGTCTCATGTAGAACTTTCTGTATTAATCTGCTATGTGGGCGT	417
OY	503	CAGCACCGCGCTGTGCTGACAGGCTTCAGGGCGAGCTGGGATGTGCTTGTCACTTCTTCAG	562
Db	418	CGGACCGCGCGCAGGCTGTGAGGCTCTCGGGCGGCTGGATGTGCTTTTGC CGGCTTTCAG	477
OY	563	ACCCCTTTCACGTGGCTTTTGTGAAGAAACAGCTGTGTCAAGGCTTTTCTGTCAAGCTTG	622
Db	478	AATCTTTAACCGCTCGGCTTCGTGGAGAAAGATGTGATTCAGGCTTTTCTGTCAAGCTTC	537
OY	623	TTAACACAGCGCTCATTTATCTGTGGAACAGATTAATA	661
Db	538	TTTGCAAGCAGCATTTTTTATCTGTGAAAGGTTATAA	576

RESULT	8
C0799635	
LOCUS	758 bp mRNA linear EST 05-AUG-2004
DEFINITION	AIRCOURT_30246635 NIH_MGC_256 mus musculus cDNA clone
IMAGE:	I09J7853 S , mRNA sequence.
ACCESSION	C0799635
VERSION	C0799635.1 GI:50987815

**KEYWORDS**  
EST.  
Mus musculus (house mouse)

**SOURCE**  
Mus musculus

**ORGANISM**  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE**  
1 (bases 1 to 758)  
NIH-MGC <http://mgc.nci.nih.gov/>.

**AUTHORS**  
National Institutes of Health, Mammalian Gene Collection (MGC)

**TITLE**  
Unpublished (1999)

**JOURNAL**  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
CDNA Library Preparation: Express Genomics  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
Plate: NDAM1195 row: n column: 06  
High quality sequence stop: 623.  
Location/Qualifiers  
1..758

**FEATURES**  
source  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30937853"  
/lab\_host="DH10B TONa"  
/clone\_1lb="NIH\_MGC\_256"  
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; cDNA was primed using oligo-dT primer: 5'-pGACTGTTCTAGATCGGAGCGCGCCCTT)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb resulted in an average insert size of 1.2 kb. This is a primary library (normalized primary library is NIH\_MGC\_257) and was constructed by Express Genomics (Frederick, MD).  
Note: this is a NIH\_MGC library"

**ORIGIN**

Query Match 24.9%; Score 220.8; DB 7; Length 758;  
Best Local Similarity 64.6%; Pred. No. 1.6e-43;  
Matches 371; Conservative 0; Mismatches 182; Indels 21; Gaps 2;

86 GACCCGCTGCGGAGCGGACGAGCTGTGCTGCGCGCACTACCTGGGGTACTGCGCCGG 145  
51 GACCCACTGTCATGAAACGACATGACGCGCTGCTGTGACTACATATTTCTTCTGCGCGCA 110  
146 GAACCCGCGACCCCGGAGCGCGCGCATTCACGCGCGCGCGCGCTGCGCGCTCGCGG 205  
111 GAGCGGACACCCGAGGACCGCCGACGCTGTGCGAGGCGCGCTTGTGCTGTG 170  
206 GCCGCGAGTTACCGGACGATTCACCGGATCTTTTCTCCGCTACTGCGGCTACCCCGGG 265  
171 ACTAGGAGATTCAGGAGGACCAAGAAATTTTCTCTCTCTCTGGAAGCGCGGGC 230  
266 AACCGTTGAGCTGTGCGCGCTGATGCGGATTCGCTCTCTCGACAGCGCGCGCGCC 325  
231 AATCGCTGAGGCTGTGAAACAGATGACATGATGATGCTCTCCAAAGACCAAGACTTC 290  
326 ACCGCGGCGAGAGTGAGCGCTGTGACCTTCCAGGAGAGCGCTCGAGAGAGCGCG 385  
291 AGCTGGAGCCAACTGGTATGCTCTGCGCTTCCGGGAGCGCTTAATCAAGGCGCT 350  
386 CTGGTGAACCGCCCGGTGAAGATGGGGCTTCCAGCGCGGCTAAAGAGCAGAGGGC 445  
351 TACATGCTGTCAAGCAGAAAT-----GGATCTGGGAATCTGTCTC 392  
446 GACGTGCGCGGAGCTGCAAGCGCGCTGTGCTGTGCTGAGCTCGCGCTCATGGG--- 502  
393 ATAGTACCCGAGACTGCTGTCTCATAGTGAATCTTGTATTAATCTGCTCATGGGCGCT 452

**RESULT 9**  
CO806531  
LOCUS  
DEFINITION  
AGENCOURT 30256704 NIH\_MGC\_256 Mus musculus cDNA clone  
IMAGE:30938765 5', mRNA sequence.  
CO806531  
CO806531.1 GI:51021322

**ACCESSION**  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

**REFERENCE**  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
CDNA Library Preparation: Express Genomics  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
Plate: NDAM1198 row: d column: 06  
High quality sequence start: 19  
High quality sequence stop: 527.  
Location/Qualifiers  
1..821

**FEATURES**  
source  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30938765"  
/lab\_host="DH10B TONa"  
/clone\_1lb="NIH\_MGC\_256"  
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; cDNA was primed using oligo-dT primer: 5'-pGACTGTTCTAGATCGGAGCGCGCCCTT)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb resulted in an average insert size of 1.2 kb. This is a primary library (normalized primary library is NIH\_MGC\_257) and was constructed by Express Genomics (Frederick, MD).  
Note: this is a NIH\_MGC library"

**ORIGIN**

Query Match 24.1%; Score 213.6; DB 7; Length 821;  
Best Local Similarity 61.3%; Pred. No. 9.7e-42;  
Matches 404; Conservative 0; Mismatches 229; Indels 26; Gaps 3;

81 TGGCGGACCCGCTGCGGAGCGGACGAGCTGTGCTGCGCGCACTACCTGGGGTACTGCG 140  
3 TGGCTTAAGACAGTTCGCGCGGAGATCCCGGATTTGCTGTGACTACATATTTCTTGGC 62  
141 CCCGGAAACCCCGGACCCCGGAGCGCGCATTCACGCCGAGGCGCGCGTGTGGCT 200  
63 CAGCGAGCGGAGACCCGAGAGCCACGCGCATGCTGTCTGAGGCGCGCTTGGCTT 122

QY 201 CCGCGGCGCCGAGGTTACGGAGATTACCGGTCCTTTTCTCCGCTACCTCGGCTACC 260  
DB 123 CTGTGATCAGGAGATCCAGAGAGACCAAGATTTTTCCTTCCTTCGCAAAAGCC 182  
QY 261 CCGGAGACCGCTTCAGCTGTGTGGCGGTGANTGGCGAATTCGTCTTCGACAGCCCCG 320  
DB 183 GGGGCAATCGCTGTGAGCTGGTGAACAAGATGGCAGATAGTTGCTTCGAAAGACCAAG 242  
QY 321 GCCCAGCTGGGGGAGATGTGTGAGCTGTGACCTTTGGCAGGACCGTTCGTGAGAGAG 380  
DB 243 ACTTGAAGTGAAGCAACTGTGTGATGCTTCGCGCTTCGCGGAGACGCTTATGATCAAG 302  
QY 381 GAGCGCTGTGATGACCGCCGCGTGAAGAGATGGGGCTTCGACCGCGGCTAAAGAGACAG 440  
DB 303 GCCCTTAAATAGGCTGTCAAGCAGAAAG-----GGATATCGGGGAATC 344  
QY 441 AGGGGAGCGTGGCGCCGGGAGCTGCGAGCGCTGTGGCTTCGAGCTCGCGGCTCATGG 500  
DB 345 GTGTCAATAGTGAAGCCGAGACTGTGTCTCATAGTGAACCTTCTGTATATCTGCTCATGG 404  
QY 501 GG----CAGCAGCGCGCTGTGAGCTGTGAGGCGGCTGGAGATGGCTTTTGTCACTTCT 557  
DB 405 GAGCGTGGCAGCGCGCCAGGCTGAGGCTCTCGCGGCTGGAGTGGCTTTTGTCCGCTTCT 464  
QY 558 TCAGAGACCCCTTTTCAGCTGTGTGTGAGAGAAACAGCTGGTCCAGGCTTTTGTGTAT 617  
DB 465 TCAGAAATCTTTTACCGCTCGGCTTCGTGAGAAAGATGTGTATTCAGGCTTTTCTGTAG 524  
QY 618 GCTTGTAAACAACAGCTTTCATTATCTGTGACACGATTAATATAGTTTAAACCTTT 677  
DB 525 GCTTCTTTTGAACAGCCCATCTTTTATCTGGAACCACTTATTA-----GTTTAAATTT 579  
QY 678 TAAACCGCTTCTACCTGCGCAACTGTGACCAACTAATAGACAGATGTGTGAGAACAA 736  
DB 580 TAAAGCGATCTTACCTACCTACTGTGAACCTCTTAAGGAACTTCTGGGANTANTNNA 638

RESULT 10  
LOCUS CA560564 524 bp mRNA linear EST 19-NOV-2002  
DEFINITION K0274C06-5N NIA Mouse Unfertilized Egg cDNA Library (Long) Mus  
CA560564  
VERSION musculus cDNA clone NIA:K0274C06 IMAGE:30051773 5', mRNA sequence.  
KEYWORDS CA560564.1 GI:25105219  
SOURCE EST.  
ORGANISM Mus musculus (house mouse)  
MUS musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 524)  
Piao Y., Kargul G.J., Dudekula, D.B., Qian Y., Luo A., Steggy, C.A.,  
Martin, P., Abda, K., Tanaka, T. and Ko, M.S.H.  
Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library  
(Long)  
Unpublished (2001)  
Other ESTs: K0274C06-3  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
Plate: K0274 row: C column: 06  
Seq primer: M13 Reverse  
High quality sequence stop: 524  
POLYA=No.

FEATURES  
source

1..524  
/location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="na:EST:K0274C06-5N"  
/db\_xref="taxon:10090"

/clone="NIA:K0274C06 IMAGE:30051773"  
/issue\_type="Unfertilized Egg"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse Unfertilized Egg cDNA Library  
(Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://igsun.grc.nia.nih.gov/cdna). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were  
extracted from a pool of 148 unfertilized eggs.  
Double-stranded cDNAs were synthesized with an O190 (dT)  
primer (Invitrogen):  
5'-TGACTAGTCTAGATCGGAGCGGCGCCCTTTTCTTTTCTTTT-3',  
treated with T4 DNA polymerase, and purified by  
ethanol-precipitation. The cDNAs were ligated to  
Lone-Linker L1-SalI, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Taq polymerase (Takara) with a primer SalI-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.5 kb. The library was constructed  
by Yulan Piao (NIA)."

## ORIGIN

Query Match 23.3%; Score 206.4; DB 6; Length 524;  
Best Local Similarity 64.6%; Pred. No. 5.3e-40;  
Matches 350; Conservative 0; Mismatches 171; Indels 21; Gaps 2;

QY 116 CTGCGCCACTACTCGGGGTAATGCGCCCGGGAACCCCGGAGCCGCGCCATCC 175  
DB 1 CTGTCTGACTAATATTTCTTCTGCGCGGGAGCGGACACCCAGAGCCACCGCCACG 60  
QY 176 ACGCCGAGCGCGCGTCTCGCTCGCGCGCCGCAAGTTACGAGATTCACCGGTCC 235  
DB 61 TCTGTCAAGCGCGCTTCTCGCTCTGTATCTAGGACATCCAGCAGACACCAAGAA 120  
QY 236 TTTTCTTCGCTTACTCTGCGTACCCCGGGAACGCTTTCAGCTGTGGCGCTGATGGCG 295  
DB 121 TTTTCTTCCTCTCTCTCGGAAGCCGGGGCAATGCGCTGAGCTGTGTAACAATGAGCA 180  
QY 296 GATTCGCTCTCTCCGACAGCCCGCCGACCTGTGGGCGAAGTGTGACCTGTGACC 355  
DB 181 GATTAAGTGTCTCTCAAGACCAAGACTTCAGCTGTGAGCCAACTGTGTATGCTCTGGCC 240  
QY 356 TTGCGAGGAGCGCTGTGAGAGAGGCGCGCTGTGACCGCCCGGTGAAGAAATGGGGC 415  
DB 241 TTGCGGGAGCGCTTATGAATCAAGCGCTTACATGTGTGTCAAGCAGAAAG----- 292  
QY 416 TTCAGCGCGGCTAAAGAGCAGAGAGGCGACGTGCGCGGGACTGCGCAGCGCTGTGG 475  
DB 293 -----GGATCTGGGGAATCGTGTCAATGTGACCCGAGACTGCTGTCAATAGTG 342  
QY 476 GCTTGTGAGCTGCGCGCTCATGGG---CAGACCGCGCGTGGCTGACGCTCAGGGC 532  
DB 343 AACTTCTGTATTAATCTGTATGTGGGTGGGACCGCGCAGAGCTGAGAGGCTTCTGGC 402  
QY 533 GAGTGGAGTGGCTTTTGTCACTTCTTCAAGACCCCTTTCACATGGCTTTTGTGAGAAA 592  
DB 403 GGTGTGATGTGCTTTTGGCGCTTCTTCAAGAAATCTTACCGCTGCGGCTTCTGAGAAAG 462  
QY 593 CAGCTGTGACAGGCTTTTGTCAATGCTTTTAAACAAGCCTTCAATTAATCTGTGACA 652  
DB 463 TTGCTGATTCAGGCTTTTCTGTCAAGCTTCTTCAACAGACATCTTTTATCTGAGAA 522  
QY 653 CG 654  
DB 523 CG 524

Qy	481	GCTAGCTCGGGGCTCATTGAGGG---CAGCACCGGCGCTGGCTGACAGGCTCAGGCGGCTG	537
Db	293	TCTGTATATATCTGTCTATGAGGGCGCTCGGCAACCGGCCACAGGCTGGAAGGCTCTCGCGGCTG	352
Qy	538	GGATGCGCTTTTGTGTCATCTTCTTGAGAACCCCTTTTCCACTGGCTTTTGGAGAAAACAGCT	597
Db	353	GGATGCGCTTTTGGCGCGCTTCTTCAAGAAATCCCTTTCCCTCGGCTTCTGGAGAAATGCT	412
Qy	598	GGTCCAGGCTTTCTGTCATGCTCTGTATCAACAAGCGCTCATTTATCTCTGGACAGCATT	657
Db	413	GATTCAAGGCTTTCTGTGACAGGCTTCTTTGCAACAGCATCTTTTATATGGAACGTTT	472
Qy	658	ATTATGAGTTTAAACCTTTTAAACCCGCTTCTACCTGCCCACTGTGACCAACTAAATGA	717
Db	473	AT-----AAGTTTAAATTTTAAAGCGATCTTACCTACCTACCTGTGAACCCCTCTAAAGA	528
Qy	718	CAGATGCTG 727	
Db	529	ACATTTCTGGG 538	

  

RESULT 12			
AA005293			
LOCUS			
DEFINITION	AA005293	206 bp	mRNA
ACCESSION	U93311.1	Soares_fetal_liver_spleen	INFLS_S1 Homo sapiens cDNA
VERSION	AA005293		
KEYWORDS	AA005293.1	GI:1448755	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		

REFERENCE	AUTHORS
1. (pages 1 to 206)	Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissee, S., Dietrich, N., Dubuq, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Merdis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, M. and Marra, M.
TITLE	Generation and analysis of 280,000 human expressed sequence tags
JOURNAL	Genome Res. 6 (9), 807-828 (1996)
MEDLINE	97044478
PUBMED	8889549
COMMENT	Contact: Wilson RK Washington University School of Medicine

Fax: 314 286 1810  
 Email: east@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Insert Length: 849    Std Error: 0.00  
 Seq primer: mob.REGA+ST  
 High quality sequence stop: 126.  
 Location/Qualifiers  
     1. .206  
         /organism="Homo sapiens"  
         /mol\_type="mRNA"  
         /db\_xref="GDB:1328597"  
         /db\_xref="taxon:9606"  
         /clone="IMAGE:428828"  
         /sex="male"  
         /dev\_stage="20 week-post conception fetus"  
         /lab\_host="DH10B (ampicillin resistant)"  
         /clone\_1lb="Soares fetal liver spleen INFILS S1"  
         /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)  
         with a modified polylinker. Site 1: Pac I; Site 2: Eco RI;  
         This is a subtracted version of the original Soares fetal  
         liver spleen INFILS library. 1st strand cDNA was primed  
         with a Pac I - oligo(dT) primer [5'





Wed Jun 8 10:19:30 2005

us-10-071-174a-1.rst

Page 12

Db 415 CAGGTGTGAGGACAAGATGCAATGGCTTCTTGAGTGAAAGAA 369

Search completed: June 7, 2005, 17:48:10  
Job time : 5400.65 secs

0



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 08:53:39 ; Search time 1000.5 Seconds  
(without alignments)  
6296.021 Million cell updates/sec

Title: US-10-071-174A-37

Perfect score: 130  
Sequence: 1 tccgcctaccctcgctacc.....acctcgcaggagcgcgcgt 130

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hcg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sbs: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114.8	88.3	582	6	BD233456 Human pro
2	114.8	88.3	585	9	HS458330 Homo sapi
3	114.8	88.3	615	9	AF285092 Homo sapi
4	114.8	88.3	726	6	CQ752105 Sequence
5	114.8	88.3	887	9	AF326964 Homo sapi
6	114.8	88.3	1168	6	BD233466 Homo sapi
7	114.8	88.3	93287	9	AC023906 Homo sapi
8	88.8	68.3	214669	2	AC018903 Homo sapi
9	55.4	42.6	1074	10	AY029163 Rattus no
10	55.4	42.6	240461	2	AC111669 Rattus no
11	51	39.2	1209	10	AF102501 Mus muscu
12	51	39.2	1225	10	AF067660 Mus muscu
13	51	39.2	1257	10	BC052690 Mus muscu
14	51	39.2	168997	10	AC133947 Mus muscu
15	51	39.2	202851	10	AC115880 Mus muscu
16	37	28.5	3389	1	SAGLCTRAA Z22577 S. anti
17	36.8	28.3	3288	6	AR034064 Sequence
18	36.8	28.3	3288	6	AR097654 Sequence
19	36.8	28.3	3288	6	AR097731 Sequence

c	20	36.8	28.3	3288	6	AR102358 Sequence
c	21	36.8	28.3	3288	6	AR104868 Sequence
c	22	36.8	28.3	3288	6	AR131079 Sequence
c	23	36.8	28.3	3288	6	AR137601 Sequence
c	24	36.8	28.3	3288	6	AR453135 Sequence
c	25	36.8	28.3	3288	6	BD023245 Method fo
c	26	36.8	28.3	3288	10	MMU31993 Mus muscu
c	27	36.8	28.3	111775	10	AC018559 Mus muscu
c	28	36.8	28.3	152316	10	AC007844 Mus muscu
c	29	36.8	28.3	170323	10	AC078896 Mus muscu
c	30	36.8	28.3	209282	10	AC135105 Mus muscu
c	31	36.2	27.8	142121	9	AC036103 Homo sapi
c	32	36.2	27.8	145631	2	AC087460 Homo sapi
c	33	36.2	27.8	173660	9	AC084693 Homo sapi
c	34	36.2	27.8	175137	2	AC018670 Homo sapi
c	35	36.2	27.8	187872	2	AC067988 Homo sapi
c	36	36	27.7	301675	1	AP005027 Streptomy
c	37	35.4	27.2	868	5	CR523779 Gallus ga
c	38	35.4	27.2	1291	8	AK106045 Oryza sat
c	39	35.4	27.2	1724	6	BD016979 Starch sy
c	40	35.4	27.2	2412	6	AK755428 Sequence
c	41	35.4	27.2	2919	8	AK101978 Oryza sat
c	42	35.4	27.2	2959	8	AF419099 Oryza sat
c	43	35.4	27.2	4422	8	AY423717 Oryza sat
c	44	35.4	27.2	151668	8	AP003509 Oryza sat
c	45	34.8	26.8	1455	6	AX653652 Sequence

#### ALIGNMENTS

RESULT 1	BD233456	582 bp	DNA	linear	PAT 17-JUN-2003
LOCUS	BD233456				
DEFINITION	Human protein having hydrophobic domain and DNA encoding the same.				
ACCESSION	BD233456.1	GI:33043226			
VERSION	BD233456.1	GI:33043226			
KEYWORDS	JP 2002519016-A/2.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	Kato, S. and Kimura, T.				
TITLE	Human protein having hydrophobic domain and DNA encoding the same				
JOURNAL	Patent: JP 2002519016-A 2 02-JUN-2002;				
COMMENT	SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC				
OS	Homo sapiens (human)				
PN	JP 2002519016-A/2				
PD	02-JUN-2002				
PF	18-JUN-1999	JP 2000557267			
PI	SEISHI KATO, TOMOKO KIMURA				
PC	C12N15/09, C07K14/47, C12N1/15, C12N1/19, C12N5/10, C12N15/00, C12N5/ PC				
CC	Human protein having hydrophobic domain and DNA encoding the				
CC	same				
FT	Key	Location/Qualifiers			
FT	source	1..582			
FT		/organism="Homo sapiens (human)".			
FEATURES	source	1..582			
		/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
ORIGIN	Query Match	88.3%	Score 114.8;	DB 6;	Length 582;
	Best Local Similarity	96.9%	Pred. No. 1.7e-15;		
	Matches 127;	Conservative 0;	Mismatches 3;	Indels 1;	Gaps 1;
QY	1	TCCGCCTACCTCGACTACCCCGGACCGCTTCAGCTGCTGCGCTGATGCGCATTC	60		

Db 163 TCCGCTTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 222  
Qy 61 GTGCTCTCCGACAGCCCGGCGGACCTT-GGGAGNAGTGTGACGCTGTGACCTTCCGA 119  
Db 223 GTGCTCTCCGACAGCCCGGCGGACCTTCCGAGGAGTGTGACGCTGTGACCTTCCGA 282  
Qy 120 GGGACGCTGCT 130  
Db 283 GGGACGCTGCT 293

RESULT 2  
HSA458330 585 bp DNA linear PRI 25-APR-2002  
LOCUS Homo sapiens NRH gene for anti-apoptotic protein.  
DEFINITION AJ458330  
ACCESSION AJ458330.1 GI:20338765  
VERSION anti-apoptotic protein; NRH gene.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 Anoucheria, A., Arnaud, E., Venet, S., Lalle, P., Gony, M., Rigal, D. and  
AUTHORS Gillet, G.  
TITLE Nr1, a human homologue of Nr-13 associates with Bcl-Xs and is an  
inhibitor of apoptosis  
JOURNAL Oncogene 20 (41), 5846-5855 (2001)  
MEDLINE 21477277  
PUBMED 11593390  
REFERENCE 2 (bases 1 to 585)  
AUTHORS Gillet, G.  
TITLE Direct Submission  
JOURNAL Submitted (23-APR-2002) Gillet G., IDCP, CNRS UMR 5086, 7 PASSAGE  
DU VERORS, 69367, FRANCE  
FEATURES  
source  
location/Qualifiers  
1..585  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
1..585  
/gene="NRH"  
1..585  
/feature="NRH"  
/function="apoptosis inhibitor"  
/codon\_start=1  
/product="anti-apoptotic protein"  
/protein\_id="CAD30221.1"  
/db\_xref="GI:20338766"  
/db\_xref="GOA:O9HD36"  
/translation="MADPLRRTRELLADYLCYCARBETPRPASTPEAVLRSAA  
RLRQIHRSPFSAYIGTGNRRELVAMDSVLSDSPTGKVVTLVTPAGTLERGP  
LVTAIRKRWGQPRLEKQBGVARDQRLVALLSRLMGHRAMLQAGGMDGFCHP  
RTPPLAFWRKQLVQAFLSCLTTAFIYLTRL"

ORIGIN  
Query Match 88.3%; Score 114.8; DB 9; Length 585;  
Best Local Similarity 96.9%; Pred. No. 1.7e-15;  
Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 TCCGCTTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 60  
Db 163 TCCGCTTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 222  
Qy 61 GTGCTCTCCGACAGCCCGGCGGACCTT-GGGAGNAGTGTGACGCTGTGACCTTCCGA 119  
Db 223 GTGCTCTCCGACAGCCCGGCGGACCTTCCGAGGAGTGTGACGCTGTGACCTTCCGA 282  
Qy 120 GGGACGCTGCT 130  
Db 283 GGGACGCTGCT 293

RESULT 3  
AF285092 615 bp mRNA linear PRI 08-NOV-2001  
LOCUS Homo sapiens Bcl-2-like protein 10 mRNA, complete cds.  
DEFINITION AF285092  
ACCESSION AF285092  
VERSION AF285092.1 GI:9837265  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 Zhang, H., Holzgreve, W. and De Geyter, C.  
AUTHORS Bcl2-L-10, a novel anti-apoptotic member of the Bcl-2 family,  
blocks apoptosis in the mitochondria death pathway but not in the  
death receptor pathway  
JOURNAL Hum. Mol. Genet. 10 (21), 2329-2339 (2001)  
MEDLINE 21548034  
PUBMED 11689480  
REFERENCE 2 (bases 1 to 615)  
AUTHORS Zhang, H.H.  
TITLE Direct Submission  
JOURNAL Submitted (05-JUL-2000) University Women's Hospital, Schanzenstr  
46, Basel 4057, Switzerland  
FEATURES  
source  
location/Qualifiers  
1..615  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q21"  
/tissue="ovary"  
1..615  
/codon\_start=1  
/product="Bcl-2-like protein 10"  
/protein\_id="AA00503.1"  
/db\_xref="GI:9837266"  
/translation="MVDLRRTRELLADYLCYCARBETPRPASTPEAVLRSAA  
RLRQIHRSPFSAYIGTGNRRELVAMDSVLSDSPTGKVVTLVTPAGTLERGP  
LVTAIRKRWGQPRLEKQBGVARDQRLVALLSRLMGHRAMLQAGGMDGFCHP  
RTPPLAFWRKQLVQAFLSCLTTAFIYLTRL"

ORIGIN  
Query Match 88.3%; Score 114.8; DB 9; Length 615;  
Best Local Similarity 96.9%; Pred. No. 1.7e-15;  
Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 TCCGCTTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 60  
Db 193 TCCGCTTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 252  
Qy 61 GTGCTCTCCGACAGCCCGGCGGACCTT-GGGAGNAGTGTGACGCTGTGACCTTCCGA 119  
Db 253 GTGCTCTCCGACAGCCCGGCGGACCTTCCGAGGAGTGTGACGCTGTGACCTTCCGA 312  
Qy 120 GGGACGCTGCT 130  
Db 313 GGGACGCTGCT 323

RESULT 4  
C0752105 726 bp DNA linear PAT 03-FEB-2004  
LOCUS Sequence 38039 from Patent W002068579.  
DEFINITION C0752105  
ACCESSION C0752105.1 GI:42387450  
VERSION  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
TITLE Kites, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof  
JOURNAL Patent: WO 02068579-A 38039 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES location/Qualifiers  
source 1..726  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 88.3%; Score 114.8; DB 6; Length 726;  
Best Local Similarity 96.9%; Pred. No. 1.6e-15;  
Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 60  
Db 193 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 252  
Qy 61 GTGCTCTCCGACAGCCCGGCGCCCACT-GGGAGNAGTGTGACGCTGTGACCTTCGCA 119  
Db 253 GTGCTCTCCGACAGCCCGGCGCCCACTGTGGGAGAGTGTGACGCTGTGACCTTCGCA 312  
Qy 120 GGGACGCTGCT 130  
Db 313 GGGACGCTGCT 323  
RESULT 5  
AF326964  
LOCUS AF326964 887 bp mRNA linear PRI 01-MAY-2001  
DEFINITION Homo sapiens BCLB (BCLB) mRNA, complete cds.  
ACCESSION AF326964  
VERSION AF326964.1 GI:13898393  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 887)  
AUTHORS Ke,N., Godzik,A. and Reed,J.C.  
TITLE Bcl-2, a novel Bcl-2 family member that differentially binds and regulates Bax and Bak  
JOURNAL J. Biol. Chem. 276 (16), 12481-12484 (2001)  
MEDLINE 21201065  
PUBMED 11278245  
REFERENCE 2 (bases 1 to 887)  
AUTHORS Ke,N., Godzik,A. and Reed,J.C.  
TITLE Direct Submission  
JOURNAL Submitted (07-DEC-2000) The Burnham Institute, 10901 N. Torrey Pines Rd., La Jolla, CA 92037, USA  
FEATURES location/Qualifiers  
source 1..887  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q21"  
1..887  
/gene="BCLB"  
50..664  
/gene="BCLB"  
/note="BCL2 family protein; Bcl-2-like"  
/codon\_start=1  
/product="BCLB"  
/protein\_id="AAK48715.1"  
/db\_xref="GI:13898394"  
/translation="MVDQLRRTTMMADPLRRTLLADYLGACAREPGEPEAPSTP EAAVLRSAARLROIHRSFSAVLGYPENRVELVALMADSVLSDSPGPTGVRVTLVT FAGTLIERGPLVTARWKKMGQPRLKEQEGVARDQRLVALLSRLLGGRANVLQAO

ORIGIN GSWDFCHFFRTPEPLAFMRKOLVOAFLSCLLTAFIYLTWTRLL"  
Query Match 88.3%; Score 114.8; DB 9; Length 887;  
Best Local Similarity 96.9%; Pred. No. 1.6e-15;  
Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 60  
Db 242 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 301  
Qy 61 GTGCTCTCCGACAGCCCGGCGCCCACT-GGGAGNAGTGTGACGCTGTGACCTTCGCA 119  
Db 302 GTGCTCTCCGACAGCCCGGCGCCCACTGTGGGAGAGTGTGACGCTGTGACCTTCGCA 361  
Qy 120 GGGACGCTGCT 130  
Db 362 GGGACGCTGCT 372  
RESULT 6  
BD233466  
LOCUS BD233466 1168 bp DNA linear PAT 17-JUL-2003  
DEFINITION Human protein having hydrophobic domain and DNA encoding the same.  
ACCESSION BD233466  
VERSION BD233466.1 GI:33043236  
KEYWORDS JP 2002519016-A/12.  
SOURCE JP 2002519016-A/12.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1168)  
AUTHORS Kato,S. and Kimura,T.  
TITLE Human protein having hydrophobic domain and DNA encoding the same  
JOURNAL Patent: JP 2002519016-A 12 02-JUL-2002;  
SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC  
COMMENT OS Homo sapiens (human)  
PN JP 2002519016-A/12  
PD 02-JUL-2002  
PF 18-JUN-1999 JP 2000557267  
PI SEISHI KATO,TOMOKO KIMURA  
PC C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/PC 00  
CC Human protein having hydrophobic domain and DNA encoding the CC  
FH Key same  
FT source 1..1168  
FT location/Qualifiers  
FEATURES 1..1168  
source /organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 88.3%; Score 114.8; DB 6; Length 1168;  
Best Local Similarity 96.9%; Pred. No. 1.5e-15;  
Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 60  
Db 169 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 228  
Qy 61 GTGCTCTCCGACAGCCCGGCGCCCACT-GGGAGNAGTGTGACGCTGTGACCTTCGCA 119  
Db 229 GTGCTCTCCGACAGCCCGGCGCCCACTGTGGGAGAGTGTGACGCTGTGACCTTCGCA 288  
Qy 120 GGGACGCTGCT 130  
Db 289 GGGACGCTGCT 299

RESULT 7  
AC023906/c

LOCUS AC023906 93287 bp DNA linear PRI 06-JUL-2001  
DEFINITION Homo sapiens chromosome 15 clone CTD-2184D3 map 15q21.2, complete  
sequence.  
AC023906  
AC023906.7 GI:14595770  
HTG.  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 93287)  
1 (bases 1 to 93287)  
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,  
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,  
Pate, D. and Hood, L.  
Sequencing of human chromosome 15 D15S146-D15S117 region  
Unpublished  
2 (bases 1 to 93287)  
Rowen, L., Madan, A., Qin, S., Abbaei, N., Baradarani, L., Birditt, B.,  
Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G.,  
James, R., Kaur, A., Madan, A., Owen, M. P., Ratcliffe, A., Shaffer, T.  
and Hood, L.  
Direct Submission  
Submitted (20-FEB-2000) Multimegabase Sequencing Center, University  
of Washington, PO BOX 357730, Seattle, WA 98195, USA  
3 (bases 1 to 93287)  
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,  
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,  
Pate, D. and Hood, L.  
Direct Submission  
Submitted (04-JUL-2001) Multimegabase Sequencing Center, Institute  
for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA  
98105, USA  
4 (bases 1 to 93287)  
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,  
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,  
Pate, D. and Hood, L.  
Direct Submission  
Submitted (06-JUL-2001) Multimegabase Sequencing Center, Institute  
for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA  
98105, USA  
On Jul 4, 2001 this sequence version replaced gi:12248292.  
----- Genome Center  
Center: Multimegabase Sequencing Center  
Center code: UMWSC  
Web site: [http://chroma.mbt.washington.edu/msg\\_www](http://chroma.mbt.washington.edu/msg_www)  
Contact: [leeuwen@systemsbiology.org](mailto:leeuwen@systemsbiology.org)  
----- Summary Statistics  
Sequencing vector: pUC18; L08752  
Chemistry: Dye-terminator Big Dye; 90% of reads  
Chemistry: Dye-Primer Big Dye; 10% of reads  
Assembly program: Phrap; version 0.990399  
Note: Data from overlapping BACs AC010674 [drafting center:  
UMWSC], AC090970 [drafting center: UMWSC], and AC016824 [drafting  
center: GTC] was added for finishing  
-----  
FEATURES  
SOURCE  
Location/Qualifiers  
1..93287  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q21.2"  
/clone="CTD-2184D3"  
/clone\_lib="Cal Tech Human BAC library D"  
/note="Data from overlapping clones CTD-2650P22 AC090970,  
RP11-47K1 AC016824, and RP11-430B1 was added and the  
consensus was determined from CTD-2184D3 to the extent  
possible"  
1..8893  
/note="Overlap with CTD-2650P22 AC090970"

```

misc_feature      1..8893
                  /note="overlap with RP11-47K1 AC016824"
unsure            6078..6140
                  /note="Low quality data"
unsure            10529..10532
                  /note="Low quality data"
unsure            14722..14808
                  /note="Low quality data"
unsure            37627
                  /note="Low quality data"
unsure            71552..71593
                  /note="Low quality data"
misc_feature      82012..93287
                  /note="overlap with RP11-430B1 AC010674"
ORIGIN
Query Match      88.3%; Score 114.8; DB 9; Length 93287;
Best Local Similarity 96.9%; Pred. No. 7.3e-16;
Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Oy      1 TCGCCTACCTCGGCTACCCGGAGACCGCTTGAGCTGTGGCGCTGATGCGGATTC 60
          |||
Db      64750 TCCGCTACCTCGGCTACCCCGGGAAACCGCTTGAGCTGTGGCGCTGATGCGGATTC 64693

Oy      61 GTGCTCCGACAGCCCCGGCCCACTT-GGGAGNATGTGTGACGCTCGTACCTTGCA 119
          |||
Db      64690 GTGCTCTCCAGACAGCCCCGGCCCACTTGAGGAGATGTGTGACGCTCGTACCTTGCA 64633

Oy      120 GGGACGCTGCT 130
          |||
Db      64630 GGGACGCTGCT 64620

RESULT 8
AC018903
LOCUS      Homo sapiens chromosome 15 clone RP11-337B11 map 15q21, LOW-PASS
DEFINITION SEQUENCE SAMPLING.
ACCESSION AC018903
VERSION AC018903.2 GI:8247797
KEYWORDS HTG; HTGS PHASED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 214669)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A.,
            Madan,A., Neebitt,R., Shaffer,T. and Hood,L.
            Sequencing of human chromosome 15 D15S146-D15S117 region
            Unpublished
            2 (bases 1 to 214669)
TITLE      Journal
REFERENCE 2 (bases 1 to 214669)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
            Direct Submission
            Submitted (22-DEC-1999) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
            On Jun 4, 2000 this sequence version replaced gi:6630517.
COMMENT      -----Genome Center
            Center: Multimegabase Sequencing Center
            Center code: UWMSC
            Web site: http://chroma.mbt.washington.edu/meg_www
            Contact: leerowen@systembiology.org
            -----Summary Statistics
            Sequencing vector: pUC18; L08752
            Chemistry: Dye-terminator Big Dye; 90% of reads
            Chemistry: Dye-Primer Big Dye; 10% of reads
            Assembly program: Phrag; version 0.990399
            -----
* NOTE: This record contains 192 individual
* sequencing reads that have not been assembled into

```

\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 806: contig of 806 bp in length  
\* 807  
\* 907 906: gap of unknown length  
\* 2068: contig of 1162 bp in length  
\* 2069  
\* 2169 2168: gap of unknown length  
\* 3020: contig of 852 bp in length  
\* 3021 3120: gap of unknown length  
\* 3121 4284: contig of 1164 bp in length  
\* 4285  
\* 4385 4384: gap of unknown length  
\* 5201: contig of 817 bp in length  
\* 5202 5301: gap of unknown length  
\* 5302 6544: contig of 1243 bp in length  
\* 6545 6644: gap of unknown length  
\* 6645 7472: contig of 828 bp in length  
\* 7473 7572: gap of unknown length  
\* 7573 8756: contig of 1184 bp in length  
\* 8757 8856: gap of unknown length  
\* 8857 9658: contig of 802 bp in length  
\* 9659 9758: gap of unknown length  
\* 9759 10905: contig of 1147 bp in length  
\* 10906 11005: gap of unknown length  
\* 11006 11811: contig of 806 bp in length  
\* 11812 11911: gap of unknown length  
\* 11912 13293: contig of 1382 bp in length  
\* 13294 13393: gap of unknown length  
\* 13394 14211: contig of 818 bp in length  
\* 14212 14311: gap of unknown length  
\* 14312 15465: contig of 1154 bp in length  
\* 15466 15565: gap of unknown length  
\* 15566 16379: contig of 814 bp in length  
\* 16380 16479: gap of unknown length  
\* 16480 17624: contig of 1145 bp in length  
\* 17625 17724: gap of unknown length  
\* 17725 18578: contig of 854 bp in length  
\* 18579 18678: gap of unknown length  
\* 18679 19774: contig of 1096 bp in length  
\* 19775 19874: gap of unknown length  
\* 19875 20700: contig of 826 bp in length  
\* 20701 20800: gap of unknown length  
\* 20801 22036: contig of 1236 bp in length  
\* 22037 22136: gap of unknown length  
\* 22138 23023: contig of 887 bp in length  
\* 23024 23123: gap of unknown length  
\* 23124 24196: contig of 1073 bp in length  
\* 24197 24296: gap of unknown length  
\* 24297 25141: contig of 845 bp in length  
\* 25142 25241: gap of unknown length  
\* 25243 26385: contig of 1144 bp in length  
\* 26386 26485: gap of unknown length  
\* 26486 27325: contig of 840 bp in length  
\* 27326 27425: gap of unknown length  
\* 27426 28691: contig of 1266 bp in length  
\* 28692 28791: gap of unknown length  
\* 28792 29655: contig of 864 bp in length  
\* 29656 29755: gap of unknown length  
\* 29756 30881: contig of 1126 bp in length  
\* 30882 30981: gap of unknown length  
\* 30983 31813: contig of 832 bp in length  
\* 31814 31913: gap of unknown length  
\* 31914 33056: contig of 1143 bp in length  
\* 33057 33156: gap of unknown length  
\* 33157 34007: contig of 851 bp in length  
\* 34008 34107: gap of unknown length  
\* 34108 35320: contig of 1213 bp in length  
\* 35321 35420: gap of unknown length

35421 36207: contig of 787 bp in length  
\* 36208 36307: gap of unknown length  
\* 36309 37521: contig of 1214 bp in length  
\* 37522 37621: gap of unknown length  
\* 37622 38420: contig of 799 bp in length  
\* 38421 38520: gap of unknown length  
\* 38521 39726: contig of 1206 bp in length  
\* 39727 39826: gap of unknown length  
\* 39827 40635: contig of 809 bp in length  
\* 40636 40735: gap of unknown length  
\* 40736 41845: contig of 1110 bp in length  
\* 41846 41945: gap of unknown length  
\* 41946 42730: contig of 785 bp in length  
\* 42731 42830: gap of unknown length  
\* 42831 43917: contig of 1087 bp in length  
\* 43918 44017: gap of unknown length  
\* 44018 44825: contig of 808 bp in length  
\* 44826 44925: gap of unknown length  
\* 44926 46114: contig of 1189 bp in length  
\* 46115 46214: gap of unknown length  
\* 46215 47031: contig of 817 bp in length  
\* 47032 47131: gap of unknown length  
\* 47132 48292: contig of 1161 bp in length  
\* 48293 48392: gap of unknown length  
\* 48393 49176: contig of 784 bp in length  
\* 49177 49276: gap of unknown length  
\* 49277 50389: contig of 1113 bp in length  
\* 50390 50489: gap of unknown length  
\* 50490 51333: contig of 844 bp in length  
\* 51334 51433: gap of unknown length  
\* 51434 52765: contig of 1332 bp in length  
\* 52766 52865: gap of unknown length  
\* 52866 53710: contig of 845 bp in length  
\* 53711 53810: gap of unknown length  
\* 53811 54965: contig of 1155 bp in length  
\* 54966 55065: gap of unknown length  
\* 55066 55910: contig of 845 bp in length  
\* 55911 56010: gap of unknown length  
\* 56011 57558: contig of 1548 bp in length  
\* 57559 57658: gap of unknown length  
\* 57659 58509: contig of 851 bp in length  
\* 58510 58609: gap of unknown length  
\* 58610 59745: contig of 1136 bp in length  
\* 59746 59845: gap of unknown length  
\* 59846 60655: contig of 810 bp in length  
\* 60656 60755: gap of unknown length  
\* 60756 61976: contig of 1221 bp in length  
\* 61977 62076: gap of unknown length  
\* 62077 62896: contig of 820 bp in length  
\* 62897 62996: gap of unknown length  
\* 62997 64128: contig of 1132 bp in length  
\* 64129 64228: gap of unknown length  
\* 64229 65026: contig of 798 bp in length  
\* 65027 65126: gap of unknown length  
\* 65127 66269: contig of 1143 bp in length  
\* 66270 66369: gap of unknown length  
\* 66370 67220: contig of 851 bp in length  
\* 67221 67320: gap of unknown length  
\* 67321 68425: contig of 1105 bp in length  
\* 68426 68525: gap of unknown length  
\* 68526 69322: contig of 797 bp in length  
\* 69323 69422: gap of unknown length  
\* 69423 70568: contig of 1146 bp in length  
\* 70569 70668: gap of unknown length  
\* 70669 71547: contig of 879 bp in length  
\* 71548 71647: gap of unknown length  
\* 71648 73629: contig of 1982 bp in length  
\* 73629 73729: gap of unknown length  
\* 73730 74587: contig of 858 bp in length  
\* 74588 74687: gap of unknown length  
\* 74688 75780: contig of 1093 bp in length  
\* 75781 75880: gap of unknown length  
\* 75881 77196: contig of 1316 bp in length

Query	Match	Best Local Similarity	Score	DB 2	Length	214669			
Matches	110	Conservative	0	Mismatches	18	Indels	1	Gaps	1
QY	3	CGCTTACTTCGGGCTACGCCCCGGGAAACCGCTTCGAGCTGCTGCGCCTGATGCGGATTCCGT	62						
Db	93457	CCCTTACTTCGGGATACCCCGAAAACCGCTTCGAGCTGCTGCGCCTGATGCGGATTCCGT	93516						
QY	63	GCTTTCGAGAGCCCCGAGCCCACTT-GGAGNAGTGTGACCTCGTACCTTCGACAG	121						
Db	93517	ACTTCGAGAGAGCCGTGCGCCACTTAGGGCAAAAAGGGAGCGCTGTGACTTCCTTCAAG	93576						
QY	122	GACGCTGCT	130						
Db	93577	GACACTTCT	93585						

RESULT	9
LOCUS	AY029163
DEFINITION	AY029163 Rattus norvegicus BCL2L10 (Bcl2l10) mRNA, complete cds.
ACCESSION	AY029163
VERSION	AY029163.1 GI:13641257
KEYWORDS	.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 1074) Itoh,T., Itoh,A. and Pleasure,D. Bcl-2-related protein family gene expression during oligodendroglial differentiation J. Neurochem. 85 (6), 1500-1512 (2003)
AUTHORS	JOURNAL MEDLINE PUBMED 12787069
TITLE	2 (bases 1 to 1074) Itoh,T., Itoh,A. and Pleasure,D. Direct Submission Submitted (29-MAR-2001) Neurology Research, The Children's Hospital of Philadelphia, Abramson Research Center, Room 516 I, 3517 Civic Center Boulevard, Philadelphia, PA 19104, USA
FEATURES	location/Qualifiers
source	1..1074 /organism="Rattus norvegicus" /mol_type="mRNA" /strain="Sprague-Dawley" /db_xref="taxon:10116" 1..1074 /gene="Bcl2l10" 12..569 /gene="Bcl2l10" /note="Boo, Di'va" /codon_start=1 /product="BCL2L10" /protein_id="AKK31792.1"
CDS	

```

/db_xref="GI:13641258"
/translation="MGDPLDQTRRLTLTYILFCARAPNTPEPLPTSEVALLRSVTS
QIQOEQDLFNSPRDYQGNRLIELVTQMADELLNDEFFNGRLVMLIAFGTLNQDR
TVRRDQRRRLRLERCYVILSLVYRLTGRHSWLEAGWDGFCQFEKNPLPPEGF
WRRLRLRLISCFATAIPYIMKCL"

```

Query Match	42.6%;	Score 55.4;	DB 10;	Length 1074;
Best Local Similarity	69.8%;	Pred. No. 0.018;		
Matches	88;	Conservative	0;	Mismatches 37; Indels 1; Gaps 1;
QY	3	CGCCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGCGCTGATGCGGATTCCGT	62	
Db	176	CTCCTTCCGGAGACTACACGAGGCAACCGCTCGAGCTGTGACACAGATGCGGATGAGTT	235	
QY	63	GCTTCTCCGACAGCCCGGCCCCACTT-GGAGNAGTGTGACGCTCGTGAACCTTCGACG	121	
Db	236	GCTCTCCAAAGACCAAGATTCACATGGGGCCGCGCTGTATGCTCCTGGCCTTCGTGGG	295	
QY	122	GAGCGCT	127	
Db	296	GAGCGCT	301	

	RESULT	10
AC111669		
LOCUS	240461 bp	DNA linear
DEFINITION	Rattus norvegicus clone CH230-54110, WORKING DRAFT SEQUENCE.	5
ACCESSION	AC111669	
VERSION	AC111669.4 GI:24818904	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.	
SOURCE	Rattus norvegicus (Norway rat)	
ORGANISM	Rattus norvegicus	

REFERENCE  
AUTHORS

1 (bases 1 to 240461)

Bukaryotis; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnesstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blych, P., Brown, M., Bryant, N., Buhey, C., Burch, P., Burrell, K., Calderon, E., Cardenas, J., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., De Souza, L., Devila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C., M., Gabis, A., Gante, R., Garcia, A., Garner, T., Garza, M., Georegeogis, E., Geer, K., Gail, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idebrld, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovat, C., Kowis, C., Kraft, C., Lebow, H., Lervan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Lin, Y., London, P., Longacre, S., Lopez, J., Lorensuthwa, L., Louisseged, H., Lozado, R., J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Matinez, E., Mawhinney, S., Mcleod, M., P., McNeill, T. Z., Meenen, E., Mlloeevillavic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, U., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okunou, G., Olarnpusagoun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polndexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M., A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajd, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Woodem, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission  
Unpublished  
2 (bases 1 to 240461)  
Worley, K. C.

Direct Submission  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 240461)  
Rat Genome Sequencing Consortium.

Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 9, 2002 this sequence version replaced gi:23196052.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GNVB  
Center clone name: CH230-54110  
----- Summary Statistics  
Assembly program: Phrap, version 0.990329  
Consensus quality: 218898 bases at least Q40  
Consensus quality: 222128 bases at least Q20  
Consensus quality: 223928 bases at least Q30  
Estimated insert size: 223140; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_difft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_difft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 9792: contig of 9792 bp in length  
\* 9793 9892: gap of unknown length  
\* 9893 118024: contig of 108133 bp in length  
\* 118025 118124: gap of unknown length  
\* 118125 236958: contig of 118833 bp in length  
\* 236959 237058: gap of unknown length

237059 238398: contig of 1340 bp in length  
\* 238399 238498: gap of unknown length  
\* 238499 240461: contig of 1963 bp in length.  
Location/Qualifiers  
1..240461  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-54110"  
1..1351  
/note="wgs\_end extension  
clone\_end:Sp6"  
6797..7581  
/note="clone boundary  
clone\_end:Sp6  
site:ECORI  
end\_sequence:BH349014"  
8640..9792  
/note="wgs\_contig"  
118125..119133  
/note="wgs\_contig"

ORIGIN  
Query Match 42.6%; Score 55.4; DB 2; Length 240461;  
Best Local Similarity 69.8%; Pred. No. 0.0073;  
Matches 88; Conservative 0; Mismatches 37; Indels 1; Gaps 1;  
QY 3 CGCTTACTGGCTTACCCCGGAAACCGTTGAGCTGTGGCGCTGATGCGGATTCCT 62  
Db 21975 CTCCTTCGCGACATCACCGGCAACCGCTGTGACTGTGTACACAGATGGCGATGAGTT 22034  
QY 63 GCTCTCGACGACCCCGGCCCACTT-GGGAGNAGTGTGACGCTCGTACCTTCGACG 121  
Db 22035 GCTCTCATATACCAAGATTCACTGGGCGCCCTGTGTACTCTTGGCTTGTGGG 22094  
QY 122 GACGCT 127  
Db 22095 GACGCT 22100

RESULT 11  
AF102501 1209 bp mRNA linear ROD 21-JAN-1999  
LOCUS AF102501 Mus musculus Bcl-2 homolog (Boc) mRNA, complete cds.  
DEFINITION AF102501  
ACCESSION AF102501.1 GI:4165137  
VERSION AF102501.1  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1209)  
AUTHORS Song, Q.-Z., Kuang, Y., Dixit, V.M. and Vincenz, C.  
TITLE Boc, a novel negative regulator of cell death, interacts with  
JOURNAL Submitted (28-OCT-1998) Pathology, University of Michigan, 1301  
Catherine Road, Ann Arbor, MI 48109, USA  
FEATURES  
source  
1..1209  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6D.F1/J"  
/db\_xref="taxon:10090"  
/chromosome="9"  
/map="D9"  
/tissue type="ovary"  
1..1209  
gene

CDS	/gene="Boo" 117. .692
-----	--------------------------

Query Match	39.2%;	Score 51;	DB 10;	Length 1209;
Best Local Similarity	67.2%;	Pred. No. 0.16;		
Matches	86;	Conservative	0;	Mismatches 41;
			Indels	1;
			Gaps	1;

Db 408 GGGACGCT 415

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

FEATURES	Location/Qualifiers
source	1. .1225

polya_signal	polya_site
PLPLGFWRRLLIQAFLSGFATAIFFIWRLL	1197. .1202
/gene="D1va"	1225

## ORIGIN

Query Match	39.2%	Score 51;	DB 10;	Length 1225;
Best Local Similarity	67.2%	Pred. No. 0.16;		
Matches	86;	Conservative	0;	Mismatches 41;
				Indels 1;
				Gaps 1;

Db

434 GGACGCT 441

KEYWORDS	MGC.
SOURCE	Mus musculus (house mouse)

REMARK	COMMENT
NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	Contact: MGC help desk



CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nigrl.nih.gov](mailto:nisc_mgc@nigrl.nih.gov)

Blakesley, R.W., Bouffard, G.G., Steen, K., Brinkley, C., Brooke, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,  
Taurgeson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
Series: IRAC Plate: 112 Row: h Column: 18  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 7304926.

## FEATURES

## source

1. 1257  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="MGC:60542 IMAGE:30052580"  
/issue\_type="Egg, unfertilized, mouse"  
/clone\_id="NIA Mouse Unfertilized Egg cDNA Library  
(Long)"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1"

## gene

1. 1257  
/gene="Bcl2l110"  
/note="synonyms: Bcl, Diva"  
/db\_xref="locusID:12049"  
/db\_xref="MGI:1330841"  
145 720  
/gene="Bcl2l110"  
/codon\_start=1  
/product="Bcl2l110 protein"  
/protein\_id="AAH52690.1"  
/db\_xref="GI:30851239"  
/db\_xref="locusID:12049"  
/db\_xref="MGI:1330841"  
/translation="MAUSQDPLHGRTRRLSDYIFPCAREPDTPPEPTSYEALLNS  
VTRQIQEHQEFSSFCESRGNRLVYKQMDKLSKODFSMSQVWLLAFACITLNN  
QGPYMAVKQKDLGNRYIVTRDCLIVFLVNLIMGRHRRLBALGMDGFCRFKN  
PLPLGFRRLILQAFISGFATAIFFIMKRL"

## CDS

## misc\_feature

268 597  
/gene="Bcl2l110"  
/note="BCL; Region: BCL (B-Cell lymphoma)"  
/db\_xref="CCD:smart00337"

## ORIGIN

Query Match 39.2%; Score 51; DB 10; Length 1257;  
Best Local Similarity 67.2%; Pred. No. 0.16;  
Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 1 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCAGCTGTGCGCTGATGCGGATTC 60  
DB 1 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCAGCTGTGCGCTGATGCGGATTC 60  
QY 316 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCAGCTGTGCGCTGATGCGGATTC 375  
DB 316 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCAGCTGTGCGCTGATGCGGATTC 375  
QY 61 GTGCTTCGACAGACCCCGGCACTTG- GAGNAGTGTGACGCTGTCGACCTTGCA 119  
DB 61 GTGCTTCGACAGACCCCGGCACTTG- GAGNAGTGTGACGCTGTCGACCTTGCA 119  
QY 376 TTGCTTCGACAGACCAAGACTTCACTGAGCACTGATGCTCTGCGCTTGCGCG 435  
DB 376 TTGCTTCGACAGACCAAGACTTCACTGAGCACTGATGCTCTGCGCTTGCGCG 435  
QY 120 GGGAGCGT 127  
DB 120 GGGAGCGT 127  
QY 436 GGGAGCGT 443  
DB 436 GGGAGCGT 443

RESULT 14

AC133947  
LOCUS AC133947 168997 bp DNA linear ROD 08-SEP-2004  
DEFINITION Mus musculus chromosome 9 clone RP24-18912, complete sequence.  
ACCESSION AC133947  
VERSION AC133947.3 GI:51921460  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

On Sep 8, 2004 this sequence version replaced gi:50199129.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)  
Project Information  
Center project name: M\_BB0189102

## FEATURES

## source

Location/Qualifiers  
1. 168997  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="9"  
/clone="RP24-18912"

## ORIGIN

Query Match 39.2%; Score 51; DB 10; Length 168997;  
Best Local Similarity 67.2%; Pred. No. 0.072;  
Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 1 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCAGCTGTGCGCTGATGCGGATTC 60  
DB 1 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCAGCTGTGCGCTGATGCGGATTC 60  
QY 166583 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCAGCTGTGCGCTGATGCGGATTC 166642  
DB 166583 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCAGCTGTGCGCTGATGCGGATTC 166642  
QY 61 GTGCTTCGACAGACCCCGGCACTTG- GAGNAGTGTGACGCTGTCGACCTTGCA 119  
DB 61 GTGCTTCGACAGACCCCGGCACTTG- GAGNAGTGTGACGCTGTCGACCTTGCA 119  
QY 166643 TTGCTTCGACAGACCAAGACTTCACTGAGCACTGATGCTCTGCGCTTGCGCG 166702  
DB 166643 TTGCTTCGACAGACCAAGACTTCACTGAGCACTGATGCTCTGCGCTTGCGCG 166702  
QY 120 GGGAGCGT 127  
DB 120 GGGAGCGT 127  
QY 166703 GGGAGCGT 166710  
DB 166703 GGGAGCGT 166710

## RESULT 15

AC115880  
LOCUS AC115880 202851 bp DNA linear ROD 02-APR-2004  
DEFINITION Mus musculus chromosome 9, clone RP24-365N15, complete sequence.  
ACCESSION AC115880  
VERSION AC115880.11 GI:46063808  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)



```

repeat_region      /rpt_family="B1F"
complement(7650, .7758)
repeat_region      /rpt_family="B4"
8301, .8421
repeat_region      /rpt_family="PB1D10"
8422, .8449
repeat_region      /rpt_family="ID_B1"
9184, .9220
repeat_region      /rpt_family="(TCTCC)n"
complement(10225, .10356)
repeat_region      /rpt_family="B1_MM"
10495, .10572
repeat_region      /rpt_family="ID1_MM"
complement(10673, .10947)
repeat_region      /rpt_family="MT1F"
complement(10978, .11186)
repeat_region      /rpt_family="MTD"
complement(11187, .11580)
repeat_region      /rpt_family="ORR1A-int"
11582, .11707
repeat_region      /rpt_family="B1F"
11776, .11824
repeat_region      /rpt_family="(CA)n"
11913, .12060
repeat_region      /rpt_family="B3A"
12061, .12103
repeat_region      /rpt_family="(CA)n"
12104, .12134
repeat_region      /rpt_family="B3A"
complement(12393, .12493)
repeat_region      /rpt_family="PB1D7"
complement(12620, .12631)
repeat_region      /rpt_family="ID5"
complement(12632, .12757)
repeat_region      /rpt_family="B1F"
complement(12758, .12829)
repeat_region      /rpt_family="ID_B1"
complement(12758, .12821)
repeat_region      /rpt_family="ID5"
complement(12956, .13319)
repeat_region      /rpt_family="ORR1B2"
13921, .14047
repeat_region      /rpt_family="B1F"
15511, .15586
repeat_region      /rpt_family="B4A"
15591, .15722
repeat_region      /rpt_family="RSINE1"
15724, .15775

```

```

Query Match      39.2%; Score 51; DB 10; Length 202851;
Best Local Similarity 67.2%; Pred. No. 0.069;
Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY      1 TCCGCTACTCGCTACCGCGGAACGCTTCGAGCTGTGCGCTGATGCGGATTCC 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      30508 TCCCTCTTCGCGAAAGCGGGGCAATGCCCTGAGCTGTGAACAGATGCAATAG 30567

QY      61 GTGCTTCGACAGCCCGGCGCCACCTGG-GAGNAGTGTGACGCTCGTGAACCTTCGCA 119
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      30568 TTGCTCTCAAGACCAAGACTTCAGCTGTGAGCCAACTGTGTATGCTCCTGGCCTTCGCG 30627

QY      120 GGGACGCT 127
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      30628 GGGACGCT 30635

```

Search completed: June 7, 2005, 21:57:53  
 Job time : 1004.5 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 08:53:44 ; Search time 127.188 Seconds

(without alignments)  
6050.630 Million cell updates/sec

Title: US-10-071-174A-37

Perfect score: 130

Sequence: 1 tcgcgcaccctccgctacc.....accttcgagggagcgtcgt 130

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 16Dec04:\*

1: geneseqn1980s:\*\n2: geneseqn1990s:\*\n3: geneseqn2000s:\*\n4: geneseqn2001as:\*\n5: geneseqn2001bs:\*\n6: geneseqn2002as:\*\n7: geneseqn2002bs:\*\n8: geneseqn2003as:\*\n9: geneseqn2003bs:\*\n10: geneseqn2003cs:\*\n11: geneseqn2003ds:\*\n12: geneseqn2004as:\*\n13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114.8	88.3	582	3	AAZ90039 Hydrophob
2	114.8	88.3	887	6	AAD46683 Human Bcl
3	114.8	88.3	1168	3	AAZ90049 Hydrophob
4	100.8	77.5	548	4	ABK41913 CDNA enco
5	100.8	77.5	548	9	ADB59580 Connectiv
6	86.8	66.8	522	4	AAH47022 Human Bcl
7	73.2	56.3	874	6	ABQ44403 Oligonuc
8	73.2	56.3	874	6	ABQ44402 Oligonuc
9	62	47.7	874	6	ABQ44404 Oligonuc
10	62	47.7	874	6	ABQ44405 Oligonuc
11	51	39.2	650	10	ADD34132 Mouse mit
12	36.8	28.3	3288	2	AAAT33800 Murine in
13	36.8	28.3	3288	2	AAV27591 Mouse int
14	36.8	28.3	3288	2	AAV27591 Murine IL
15	36.8	28.3	3288	3	AAAS1987 Murine in
16	36.8	28.3	3288	3	AAAS2145 Murine in
17	36.8	28.3	3288	3	AAAS9870 Murine in
18	36.8	28.3	3288	3	AAAG1238 Murine IL
19	36.8	28.3	3288	4	AAEF5716 Murine IL
20	36.8	28.3	3288	4	AAD02813 Murine in

C 21	36.8	28.3	3288	12	ADJ88256	Adj88256 Murine IL
C 22	36.8	28.3	3288	12	ADO91243	Adq91243 Mouse CDN
C 23	36.8	28.3	3288	12	ADR01299	Adr01299 Mouse int
24	35.4	27.2	1724	6	ABA01821	AbA01821 Rice strar
25	35.4	27.2	2412	10	ADC07863	Adc07863 Rice DNA
26	35.4	27.2	2412	12	ADJ40034	Adj40034 Plant CDN
27	35.4	27.2	2607	8	ACC70868	Acc70868 Rice strar
28	35.4	27.2	2607	8	ACC70869	Acc70869 Rice strar
29	35.4	27.2	5928	8	ACC70867	Acc70867 Rice strar
30	35.4	27.2	5935	8	ACC70866	Acc70866 Rice strar
31	34.8	26.8	1455	8	ADA70199	Ada70199 Rice gene
32	33.8	26.0	24379	2	AAV25925	AAV25925 Streptomy
33	33.8	26.0	24379	2	AAV25925	AAV25925 Streptomy
34	33.6	25.8	1077	9	ADA48367	Ada48367 Rice gene
35	33.6	25.8	1077	12	ADJ39313	Adj39313 Plant CDN
C 36	33.4	25.7	1125	2	AAQ36369	Aaq36369 T.thermop
C 37	33.4	25.7	1759	10	ABZ57164	Abz57164 Human met
38	33	25.4	1911	6	ABZ78669	Abz78669 S. ghanae
39	32.6	25.1	47988	10	ADB66070	AdB66070 Streptomy
40	33.4	24.9	675	12	ADQ75857	Adq75857 Barley ly
C 41	32.2	24.8	792	4	AAH70552	Aah70552 Human cer
42	32.2	24.8	17083	10	ABZ58812	Abz58812 S. cinna
43	32	24.6	992	4	ABL12245	AbL12245 Drosophi
44	32	24.6	2859	10	ADF30786	Adf30786 Soil meta
45	32	24.6	3486	8	ABQ77425	Abq77425 Human GCD

## ALIGNMENTS

RESULT 1	
AAZ90039	standard; CDNA, 582 BP.
XX	
XX	
AAZ90039;	
XX	
DT	09-MAY-2000 (first entry)
XX	
DE	Hydrophobic domain containing protein clone HP02403 coding sequence.
XX	
XX	Hydrophobic domain; clone HP02403; nutritional supplement; SCID; HIV;
KW	cell proliferation; immune stimulant; immune deficiency; tumour; pain;
KW	rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
KW	myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
KW	anti-inflammatory; infection; bodily characteristic; ss.
OS	Homo sapiens.
XX	
PN	MO20000506-A2.
XX	
PD	06-JAN-2000.
XX	
PF	18-JUN-1999; 99WO-JP003242.
XX	
PR	26-JUN-1998; 98JP-00180008.
XX	
PA	(SAGA ) SAGAMI CHEM RES CENT.
XX	(PROT-) PROTEGENE INC.
XX	
PI	Kato S, Kimura T;
XX	
DR	WPI, 2000-160665/14.
DR	P-PSDB; AAV78802.
XX	
PT	Novel human proteins having hydrophobic domains used for research and
PT	diagnostic purposes.
XX	
PS	Claim 3; Page 84; 117pp; English.
XX	
CC	This sequence represents the hydrophobic domain containing protein, clone
CC	HP00631 coding region. The sequence is isolated from a human stomach
CC	cancer cell line. The HP02403 protein contains one putative transmembrane
CC	domain. The protein shows homology to the Japanese quail apoptosis

CC regualtor NR-13. The invention relates to human proteins with hydrophobic  
CC domains, the DNA and the cDNA encoding them. The polynucleotides and  
CC proteins are predicted to have biological activities which make them  
CC suitable for treating, preventing or ameliorating medical conditions in  
CC humans and animals. Suggested activities include nutritional activity  
CC (nutritional source or supplement); cytokine and cell  
CC proliferation/differentiation activity; immune stimulating (e.g. as  
CC vaccines) or suppressing activity (e.g. to treat various immune  
CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic  
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary  
CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin  
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease  
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and  
CC organ transplantation); haematopoiesis regulating activity (e.g. in  
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth  
CC activity (e.g. wound healing and tissue repair, ulcers, burns,  
CC periodontal disease); activin/inhibin activity; chemocactive/chemokinetic  
CC activity; haemostatic and thrombolytic activity (e.g. treating  
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and  
CC tumour inhibition activity. The polynucleotides are also stated to be  
CC useful for gene therapy. Other activities include inhibiting infections  
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,  
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,  
CC skin, effecting biorhythms or circadian cycles; enhancing fertility;  
CC treatment of depression; treatment of pain; hormonal or endocrine  
CC activity. The polynucleotides may also be used for recombinant expression  
CC of the protein

XX Sequence 582 BP; 85 A; 193 C; 191 G; 113 T; 0 U; 0 Other;

SO Query Match 88.3%; Score 114.8; DB 3; Length 582;

Best Local Similarity 96.9%; Pred. No. 1.6e-19;

Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 TCCGCTTACTCTGGGTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 60

Db 163 TCCGCTTACTCTGGGTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 222

QY 61 GTGCTCTCCGACAGCCCGGACCGCTTGGAGNAGTGATGACCTGTGACCTTCGCA 119

Db 223 GTGCTCTCCGACAGCCCGGACCGCTTGGAGNAGTGATGACCTGTGACCTTCGCA 282

QY 120 GGGACGCTGCT 130

Db 283 GGGACGCTGCT 293

RESULT 2  
AAD46683 standard; DNA; 887 BP.

XX AAD46683;

XX 27-JAN-2003 (first entry)

DT Human Bcl-B DNA.

DE Human Bcl-B DNA.

XX Human; Bcl-2; Bcl-B; therapy; apoptosis; cell degenerative disorder;

XX proliferative disorder; muscle degeneration; Alzheimer's disease; CJD;

XX Creutzfeldt-Jacob's disease; Machado-Joseph disease; MJD; transgenic;

XX Parkinson's disease; Huntington's disease; HD; spinocerebellar ataxia;

XX SCA; dentatorubropallidoluyisan atrophy; DRPLA; Kennedy's disease;

XX stroke; ischaemia; head trauma; neoplasia; anticonvulsant; vulnerrary;

XX neurotropic; neuroprotective; cytosstatic; immunosuppressive; vasotropic;

XX cerebroprotective; autoimmune disorder; chromosome 15; gene; ds.

XX Homo sapiens.

XX OS

XX Key Location/Qualifiers

XX CDS 50..664

XX FT /\*tag= a

XX FT /product= "Human Bcl-B protein"

PN WO200272601-A2.

XX 19-SEP-2002.

XX 07-FEB-2002; 2002WO-US003547.

XX 07-FEB-2001; 2001US-0267166P.

XX 07-FEB-2002; 2002US-00071174.

XX (BURN-) BURNHAM INST.

XX Reed JC, Ke N, Godzik A;

XX WPI; 2002-723312/78.

XX P-PSDB; AAE29097.

PT New isolated or recombinant Bcl-B nucleic acids and polypeptides, for

PT treating a disorder associated with apoptosis, such as cell degenerative

PT or proliferative disorder e.g. cancer, Alzheimer's disease or Parkinson's

PT disease.

XX Claim 11; Page 74; 82pp; English.

XX The invention relates to human member of Bcl-2 family Bcl-B protein and

XX its corresponding nucleic acid. Bcl-B is useful in treating a subject

XX having or at risk of a disorder associated with apoptosis, such as cell

XX degenerative or proliferative disorder like neural or muscle

XX degeneration, e.g. Alzheimer's disease, Creutzfeldt-Jacob's disease

XX (CJD), Machado-Joseph disease (MJD), Parkinson's disease, Huntington's

XX disease (HD), spinocerebellar ataxias 1, 2 and 6 (SCA-1, -2 and -6),

XX dentatorubropallidoluyisan atrophy (DRPLA), Kennedy's disease, stroke,

XX ischaemia, head trauma, neoplasia, autoimmune disorder or fibrotic

XX condition. The transgenic animals are used as in vivo models to study

XX apoptosis and potential therapies for apoptosis. The present sequence is

XX human Bcl-B DNA. Bcl-B gene is located at chromosome 15

SO Sequence 887 BP; 171 A; 270 C; 269 G; 177 T; 0 U; 0 Other;

Query Match 88.3%; Score 114.8; DB 6; Length 887;

Best Local Similarity 96.9%; Pred. No. 1.6e-19;

Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 TCCGCTTACTCTGGGTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 60

Db 242 TCCGCTTACTCTGGGTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGCGGATTCC 301

QY 61 GTGCTCTCCGACAGCCCGGACCGCTTGGAGNAGTGATGACCTGTGACCTTCGCA 119

Db 302 GTGCTCTCCGACAGCCCGGACCGCTTGGAGNAGTGATGACCTGTGACCTTCGCA 361

QY 120 GGGACGCTGCT 130

Db 362 GGGACGCTGCT 372

RESULT 3  
AAZ90049 standard; cDNA; 1168 BP.

XX AAZ90049;

XX 09-MAY-2000 (first entry)

DE Hydrophobic domain containing protein clone HP02403 nucleotide sequence.

XX Hydrophobic domain; clone HP02403; nutritional supplement; SCID; HIV;

XX cell proliferation; immune stimulant; immune deficiency; tumour; pain;

XX rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;

XX myasthenia gravis; haematopoiesis regulator; tissue growth; depression;

XX anti-inflammatory; infection; bodily characteristic; ss.

XX Homo sapiens.

XX OS

PN WO20000506-A2.  
XX 06-JAN-2000.  
PD 18-JUN-1999; 99WO-JP003242.  
XX 26-JUN-1998; 98JP-00180008.  
XX (SAGA ) SAGAMI CHEM RES CENT.  
PA (PROT-) PROTEGENE INC.  
XX Kato S, Kimura T;  
XX WPI: 2000-160665/14.  
DR P-PSDB: AAV78802.  
XX  
PT Novel human proteins having hydrophobic domains used for research and  
PT diagnostic purposes.  
XX  
XX Claim 4; Page 92-94; 117pp; English.  
XX  
XX This sequence represents the hydrophobic domain containing protein, clone  
CC HP00631 nucleotide sequence. The sequence is isolated from a human  
CC stomach cancer cell line. The HP02403 protein contains one putative  
CC transmembrane domain. The protein shows homology to the Japanese quail  
CC apoptosis regulator NR-13. The invention relates to human proteins with  
CC hydrophobic domains, the DNA and the cDNA encoding them. The  
CC polynucleotides and proteins are predicted to have biological activities  
CC which make them suitable for treating, preventing or ameliorating medical  
CC conditions in humans and animals. Suggested activities include  
CC nutritional activity (nutritional source or supplement); cytokine and  
CC cell proliferation/differentiation activity; immune stimulating (e.g. as  
CC vaccines) or suppressing activity (e.g. to treat various immune  
CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic  
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary  
CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin  
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease  
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and  
CC organ transplantation); haematopoiesis regulating activity (e.g. in  
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth  
CC activity (e.g. wound healing and tissue repair, ulcers, burns,  
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic  
CC activity; haemostatic and thrombolytic activity (e.g. treating  
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and  
CC tumour inhibition activity. The polynucleotides are also stated to be  
CC useful for gene therapy. Other activities include inhibiting infections  
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,  
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,  
CC skin, effecting biohythms or cardiac cycles; enhancing fertility;  
CC treatment of depression; treatment of pain; hormonal or endocrine  
CC activity. The polynucleotides may also be used for recombinant expression  
CC of the protein.  
XX  
XX Sequence 1168 BP; 276 A; 312 C; 300 G; 280 T; 0 U; 0 Other;  
QY  
QY Query Match 88.3%; Score 114.8; DB 3; Length 1168;  
QY Best Local Similarity 96.9%; Pred. No. 1.6e-19;  
Db Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 1 TCCGCTCACTCGGTACCCCGGAAACGCTTCGAGCTGTGGCCGTGATGCC 60  
Db 169 TCCGCTCACTCGGTACCCCGGAAACGCTTCGAGCTGTGGCCGTGATGCC 228  
QY 61 GTGCTCTCGACAGCCCGGCGCCCACT-GGAGNAGTGTGACGCTGTGACCTTGC 119  
Db 229 GTGCTCTCGACAGCCCGGCGCCCACTGTGGGCAAGTGTGACGCTGTGACCTTGC 288  
QY 120 GGGACGCTGCT 130  
Db 289 GGGACGCTGCT 299

RESULT 4

ABK41913  
ID ABK41913 standard; cDNA; 548 BP.  
XX  
XX AC ABK41913;  
XX DT 21-MAY-2002 (first entry)  
XX  
XX DE cDNA encoding novel human connective tissue related polypeptide #301.  
XX KW Human; connective tissue related disorder; cancer; gene therapy;  
XX KW cytoskeletal; gene; ss.  
XX OS Homo sapiens.  
XX  
XX PN WO20015343-A1.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US001322.  
XX  
XX PR 31-JAN-2000; 2000US-0179065P.  
XX PR 04-FEB-2000; 2000US-0180628P.  
XX PR 24-FEB-2000; 2000US-0184664P.  
XX PR 02-MAR-2000; 2000US-0186350P.  
XX PR 16-MAR-2000; 2000US-0189874P.  
XX PR 17-MAR-2000; 2000US-0190076P.  
XX PR 18-APR-2000; 2000US-0198123P.  
XX PR 19-MAY-2000; 2000US-020515P.  
XX PR 07-JUN-2000; 2000US-0209467P.  
XX PR 28-JUN-2000; 2000US-0214886P.  
XX PR 30-JUN-2000; 2000US-0215135P.  
XX PR 07-JUL-2000; 2000US-0216647P.  
XX PR 07-JUL-2000; 2000US-0216880P.  
XX PR 11-JUL-2000; 2000US-0217487P.  
XX PR 11-JUL-2000; 2000US-0217496P.  
XX PR 14-JUL-2000; 2000US-0218290P.  
XX PR 26-JUL-2000; 2000US-0220963P.  
XX PR 26-JUL-2000; 2000US-0220964P.  
XX PR 14-AUG-2000; 2000US-0224518P.  
XX PR 14-AUG-2000; 2000US-0224519P.  
XX PR 14-AUG-2000; 2000US-0225213P.  
XX PR 14-AUG-2000; 2000US-0225214P.  
XX PR 14-AUG-2000; 2000US-0225266P.  
XX PR 14-AUG-2000; 2000US-0225267P.  
XX PR 14-AUG-2000; 2000US-0225268P.  
XX PR 14-AUG-2000; 2000US-0225270P.  
XX PR 14-AUG-2000; 2000US-0225447P.  
XX PR 14-AUG-2000; 2000US-0225757P.  
XX PR 14-AUG-2000; 2000US-0225758P.  
XX PR 14-AUG-2000; 2000US-0225759P.  
XX PR 18-AUG-2000; 2000US-0226279P.  
XX PR 22-AUG-2000; 2000US-0226681P.  
XX PR 22-AUG-2000; 2000US-0226868P.  
XX PR 22-AUG-2000; 2000US-0227182P.  
XX PR 23-AUG-2000; 2000US-0227009P.  
XX PR 30-AUG-2000; 2000US-0228924P.  
XX PR 01-SEP-2000; 2000US-0229287P.  
XX PR 01-SEP-2000; 2000US-0229343P.  
XX PR 01-SEP-2000; 2000US-0229344P.  
XX PR 01-SEP-2000; 2000US-0229345P.  
XX PR 05-SEP-2000; 2000US-0229509P.  
XX PR 05-SEP-2000; 2000US-0229513P.  
XX PR 06-SEP-2000; 2000US-0230437P.  
XX PR 06-SEP-2000; 2000US-0230438P.  
XX PR 08-SEP-2000; 2000US-0231242P.  
XX PR 08-SEP-2000; 2000US-0231243P.  
XX PR 08-SEP-2000; 2000US-0231244P.  
XX PR 08-SEP-2000; 2000US-0231413P.  
XX PR 08-SEP-2000; 2000US-0231414P.  
XX PR 08-SEP-2000; 2000US-0232080P.  
XX PR 12-SEP-2000; 2000US-0231968P.  
XX PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0233398P.  
 PR 14-SEP-2000; 2000US-0233399P.  
 PR 14-SEP-2000; 2000US-0233400P.  
 PR 14-SEP-2000; 2000US-0233401P.  
 PR 14-SEP-2000; 2000US-0233402P.  
 PR 14-SEP-2000; 2000US-0233403P.  
 PR 14-SEP-2000; 2000US-0233404P.  
 PR 14-SEP-2000; 2000US-0233405P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234224P.  
 PR 25-SEP-2000; 2000US-0234977P.  
 PR 25-SEP-2000; 2000US-0234978P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235835P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239335P.  
 PR 20-OCT-2000; 2000US-0239337P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251883P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-565190/63.  
 DR P-PSDB; AAU86735.  
 XX  
 XX  
 PT Nucleic acid encoding novel connective tissue associated polypeptides,  
 PT used in diagnosing, preventing, treating or ameliorating a disorder such  
 PT as cancer or rheumatoid arthritis.  
 XX  
 PS Claim 4; SEQ ID NO 311; 673bp; English.  
 XX  
 CC The present invention relates to the isolation of novel human connective  
 CC tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide  
 CC (cDNA and genomic) sequences encoding them. The sequences of the  
 CC invention are useful in the diagnosis, treatment, prevention and/or  
 CC prognosis of diseases associated with connective tissue(s), including  
 CC cancer. The polynucleotide sequences of the invention are also useful in  
 CC gene therapy. ABX41613-ABX42101 represent cDNA sequences encoding the  
 CC novel human connective tissue related polypeptides. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_ptc\_sequences  
 XX  
 XX Sequence 548 BP; 99 A; 154 C; 167 G; 121 T; 0 U; 7 Other;  
 SQ  
 Query Match 77.5%; Score 100.8; DB 4; Length 548;  
 Best Local Similarity 92.9%; Pred. No. 5e-16; 8; Indels 0; Gaps 0;  
 Matches 105; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 18 CCGCGGAACCGCTTGCAGTGTGGCGCTGATGCGGATTCGTCTTCCGACAGCCC 77  
 Db 61 CCACGCGTCGCGCTTGCAGTGTGGCGCTGATGCGGATTCGTCTTCCGACAGCCC 120  
 QY 78 CCGCGCCACCTTGAGAGNAGTGTGACGCTGCTGACCTTCCGAGGAGCGCTGCT 130  
 Db 121 CCGCGCCACCTTGAGAGTGTGACGCTGCTGACCTTCCGAGGAGCGCTGCT 173  
 RESULT 5  
 ADB59580  
 ID ADB59580 standard; cDNA; 548 BP.  
 XX  
 AC ADB59580;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Connective tissue related polynucleotide #301.  
 XX  
 KW cytosolic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;  
 KW antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;  
 KW antiinflammatory; antiallergic; antilastmatic; dermatological;  
 KW nephrotoxic; virucide; fungicide; antibacterial; antiparasitic;  
 KW gene therapy; ds; connective tissues disorder; rheumatoid arthritis;  
 KW systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;  
 KW cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disease;  
 KW atherosclerosis; myocarditis; cardiopulmonary bypass complication;  
 KW autoimmune disease; multiple sclerosis; allergic reaction; asthma;  
 KW rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;  
 KW gastrointestinal disorder; inflammatory bowel disease;



KW organ transplant rejection; immune system disorder; Bruton's disease;  
KW X-linked lymphoproliferative syndrome;  
KW B-cell lymphoproliferative disorder; HIV; AIDS; infection;  
KW chromosome identification; chromosome mapping;  
KW connective tissue related polynucleotide; gene; ss.  
OS Homo sapiens.  
XX US2003054375-A1.  
PN 20-MAR-2003.  
XX 07-MAR-2002; 2002US-00092154.  
PF 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-020515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226861P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0234000P.  
PR 14-SEP-2000; 2000US-0234011P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 14-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250191P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764847.  
XX  
XX (HUMA-) HUMAN GENOME SCT INC.  
XX  
XX  
XX Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI: 2003-634869/60.  
XX P-PSDB; ADB60069.  
XX  
XX New connective tissue-related polypeptides and polynucleotides, useful  
XX for treating, preventing and/or prognosing e.g. disorders of connective  
XX tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or  
XX neoplasias.  
XX  
XX Claim 1; SEQ ID NO 311; 248bp; English.  
XX  
XX The invention describes an isolated nucleic acid molecule (1), which  
XX comprises a sequence that is at least 95 % identical to a connective  
XX tissue-related polynucleotide encoding connective tissue antigens (CTA).  
XX The polypeptide or polynucleotide is useful for preventing, treating, or  
XX ameliorating medical conditions in a mammal. The connective tissue  
XX polypeptides, polynucleotides and antibodies are particularly useful for  
XX treating, preventing and/or prognosing disorders of connective tissues  
XX (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,  
XX scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or  
XX neoplasias (e.g. leukemia), neurodegenerative disorders (e.g.  
XX Alzheimer's disease, or Parkinson's disease), cardiovascular diseases  
XX (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass  
XX complications), autoimmune diseases (e.g. systemic lupus erythematosus,  
XX rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.  
XX  
XX Query Match 77.5%; Score 100.8; DB 9; Length 548;  
XX Best Local Similarity 92.9%; Pred. No. 5e-16;  
XX Matches 105; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
XX  
XX 18 CCCCCGAAACGCTTCGAGCTGTGGCGCTGATGCGGATTCCTCCGACAGCC 77  
XX Db 61 CACCGCTCGGCTTCGAGCTGTGGCGCTGATGCGGATTCCTCCGACAGCC 120  
XX  
XX 78 CGGCCCCACCTGGAGNAGTGTGACGCTCTGACCTTCGACGAGGACGCTGCT 130  
XX Db 121 CGGCCCCACCTGGAGNAGTGTGACGCTCTGACCTTCGACGAGGACGCTGCT 173  
XX  
XX RESULT 6  
XX AAH47022  
XX ID AAH47022 standard; cDNA; 522 BP.  
XX AC  
XX AAH47022;  
XX  
XX 29-OCT-2001 (first entry)  
XX  
XX Human Bcl-2-like polypeptide encoding cDNA (clone HL1BE40).  
XX  
XX Bcl-2-like polypeptide; autoimmune disorder; allergy; immunomodulatory;  
XX respiratory; cardiovascular; antiarthritic; immunostimulant; vaccine;  
XX immunosuppressive; antiinflammatory; gene therapy; ss.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 44..448  
XX FT /\*tag= a  
XX FT /product= "Bcl-2-like polypeptide"  
XX PN WO200157060-A1.  
XX XX  
XX 09-AUG-2001.

XX  
XX 31-JAN-2001; 2001WO-US003080.  
XX  
XX  
XX 01-FEB-2000; 2000US-0179487P.  
XX 07-FEB-2000; 2000US-0180697P.  
XX  
XX  
XX (HUMA-) HUMAN GENOME SCT INC.  
XX  
XX  
XX Ruben SM, Duan DR, N1 J;  
XX  
XX WPI: 2001-476279/51.  
XX P-PSDB; AAB85666.  
XX  
XX  
XX Nucleic acids encoding human Bcl-2-like polypeptides, useful for  
XX preventing, diagnosing and/or treating.  
XX  
XX Claim 1; Page 276; 285pp; English.  
XX  
XX The invention provides nucleic acid molecules (NAM1) encoding 4 human Bcl  
XX -2-like polypeptides (PEP1). The NAM1 and PEP1 may be used in the  
XX prevention, diagnosis and treatment of diseases associated with  
XX inappropriate Bcl-2-like polypeptides' expression. The NAM1 may be used  
XX to produce the soluble Bcl-2-like polypeptides by standard recombinant  
XX methodology. The polypeptides may also be used as antigens in the  
XX production of antibodies against Bcl-2 and in assays to identify  
XX modulators of Bcl-2 expression and activity. The anti-Bcl-2 antibodies  
XX and antagonists may be used to down regulate expression and activity. The  
XX anti-PEP1 antibodies may also be used as diagnostic agents for detecting  
XX the presence of Bcl-2 polyps in samples (e.g. by enzyme linked  
XX immunosorbent assay (ELISA)). Disorders that may be prevented, diagnosed  
XX and/or treated by the above methods include, immunodeficiencies (e.g. a  
XX gammaglobulinemia and B cell lymphoproliferative disorder), autoimmune  
XX disorders (e.g. rheumatoid arthritis and Grave's disease), allergic  
XX reactions, inflammations, respiratory diseases and cardiovascular  
XX disorders (a full list of disorders is given in the specification). The  
XX present sequence represents a human Bcl-2-like polypeptide encoding cDNA  
XX  
XX Sequence 522 BP; 102 A; 148 C; 154 G; 118 T; 0 U; 0 Other;  
XX  
XX Query Match 66.8%; Score 86.8; DB 4; Length 522;  
XX Best Local Similarity 96.1%; Pred. No. 1.6e-12;  
XX Matches 99; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
XX  
XX 29 GCTTCGAGCTGTGGCGCTGATGCGGATTCCTGCTCTCCGACAGCCCGGACCACT 88  
XX Db 1 GCTTCGAGCTGTGGCGCTGATGCGGATTCCTGCTCTCCGACAGCCCGGACCACT 60  
XX  
XX 89 -GGAGNAGTGTGACGCTCTGACCTTCGACGAGGACGCTGCT 130  
XX Db 61 GGGGACAGAGTGTGACGCTCTGACCTTCGACGAGGACGCTGCT 103  
XX  
XX RESULT 7  
XX ABQ44403  
XX ID ABQ44403 standard; DNA; 874 BP.  
XX AC  
XX ABQ44403;  
XX  
XX 12-JUL-2002 (first entry)  
XX  
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 30994.  
XX  
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
XX drug; side effect; cancer; central nervous system; cardiovascular;  
XX gastrointestinal; respiratory system; single nucleotide polymorphism;  
XX SNP; cell differentiation; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200218632-A2.  
XX PN  
XX 07-MAR-2002.  
XX PD  
XX

PF 01-SEP-2001; 2001WO-EP010074.  
XX  
XX 01-SEP-2000; 2000DE-01043826.  
PR 05-SEP-2000; 2000DE-01044543.  
RR  
XX  
XX (EPiG-) EPIGENOMICS AG.  
PA  
XX  
XX Olek A, Piepenbrock C, Berlin K, Guetig D;  
PI WPI; 2002-371829/40.  
DR  
XX  
XX Determining the degree of cytosine methylation in genomic DNA, useful for  
PT diagnosis and prognosis, comprises selective hybridization of amplicons  
PT from chemically treated DNA.  
XX  
XX  
PS Claim 12; 56pp + Sequence Listing; 56pp; German.  
XX  
XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one member,  
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
CC degree of hybridisation to both classes is determined from the label on  
CC the amplicon. From the ratio of labels hybridised to the two classes of  
CC oligomers, the degree of methylation is calculated. The method is used:  
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
CC and of a wide range of diseases, e.g. cancer, disorders of the central  
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
CC particularly by detecting mutations or single nucleotide polymorphisms  
CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
CC investigating cell differentiation. The method allows the methylation  
CC status of many C residues to be determined simultaneously. ABQ13410-  
CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
CC for determining the degree of cytosine methylation described in the  
CC disclosure of the invention  
XX  
SQ Sequence 874 BP; 374 A; 291 C; 90 G; 119 T; 0 U; 0 Other;  
Query Match 56.3%; Score 73.2; DB 6; Length 874;  
Best Local Similarity 77.1%; Pred. No. 4.1e-09;  
Matches 101; Conservative 0; Mismatches 29; Indels 1; Gaps 1;  
QY 1 TCCGCTTACCTCGGCTACCCGGGAACCGCTTCGAGCTGGCGCTGATGCGGATTCC 60  
Db 341 TCCGCTTACCTCGGCTACCCGGGAACCGCTTCGAGCTTCGAACTAATACGAAATTC 400  
QY 61 GTGCTCTCCGACAGCCCGGCGCCGACCTGGGA-GNAGTGTGAGCGCTGTGACCTTGGCA 119  
Db 401 GTACTCTCCGACACCCCGGACCCGACCTTAACAAATATATACCTGTAACTTGGCA 460  
QY 120 GGGACGCTGCT 130  
Db 461 AAAACGCTACT 471  
RESULT 8  
ABQ44402/C  
ID ABQ44402 standard; DNA; 874 BP.  
XX  
XX  
AC ABQ44402;  
DT 12-JUL-2002 (first entry)  
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30993.  
XX  
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KM drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
XX SNP; cell differentiation; ds.  
XX  
OS Homo sapiens.

XX  
XX WO200218632-A2.  
PN  
XX  
XX 07-MAR-2002.  
PD  
XX  
XX 01-SEP-2001; 2001WO-EP010074.  
PF  
XX  
XX 01-SEP-2000; 2000DE-01043826.  
PR 05-SEP-2000; 2000DE-01044543.  
RR  
XX  
XX (EPiG-) EPIGENOMICS AG.  
PA  
XX  
XX Olek A, Piepenbrock C, Berlin K, Guetig D;  
PI WPI; 2002-371829/40.  
DR  
XX  
XX Determining the degree of cytosine methylation in genomic DNA, useful for  
PT diagnosis and prognosis, comprises selective hybridization of amplicons  
PT from chemically treated DNA.  
XX  
XX  
PS Claim 12; 56pp + Sequence Listing; 56pp; German.  
XX  
XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one member,  
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
CC degree of hybridisation to both classes is determined from the label on  
CC the amplicon. From the ratio of labels hybridised to the two classes of  
CC oligomers, the degree of methylation is calculated. The method is used:  
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
CC and of a wide range of diseases, e.g. cancer, disorders of the central  
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
CC particularly by detecting mutations or single nucleotide polymorphisms  
CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
CC investigating cell differentiation. The method allows the methylation  
CC status of many C residues to be determined simultaneously. ABQ13410-  
CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
CC for determining the degree of cytosine methylation described in the  
CC disclosure of the invention  
XX  
SQ Sequence 874 BP; 119 A; 90 C; 291 G; 374 T; 0 U; 0 Other;  
Query Match 56.3%; Score 73.2; DB 6; Length 874;  
Best Local Similarity 77.1%; Pred. No. 4.1e-09;  
Matches 101; Conservative 0; Mismatches 29; Indels 1; Gaps 1;  
QY 1 TCCGCTTACCTCGGCTACCCGGGAACCGCTTCGAGCTGGCGCTGATGCGGATTCC 60  
Db 534 TCCGCTTACCTCGGCTACCCGGGAACCGCTTCGAACTAATACGAAATTC 475  
QY 61 GTGCTCTCCGACAGCCCGGCGCCGACCTGGGA-GNAGTGTGAGCGCTGTGACCTTGGCA 119  
Db 474 GTACTCTCCGACACCCCGGACCCGACCTTAACAAATATATACCTGTAACTTGGCA 415  
QY 120 GGGACGCTGCT 130  
Db 414 AAAACGCTACT 404  
RESULT 9  
ABQ44404  
ID ABQ44404 standard; DNA; 874 BP.  
XX  
XX  
AC ABQ44404;  
DT 12-JUL-2002 (first entry)  
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30995.  
XX  
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KM

KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200218632-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PF 01-SEP-2001; 2001WO-EP010074.  
 XX  
 PR 01-SEP-2000; 2000DE-01043826.  
 PR 05-SEP-2000; 2000DE-01044543.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;  
 DR WPI; 2002-371829/40.  
 XX  
 PT Determining the degree of cytosine methylation in genomic DNA, useful for  
 PT diagnosis and prognosis, comprises selective hybridization of amplicons  
 PT from chemically treated DNA.  
 XX  
 PS Claim 12; 56bp + Sequence Listing; 56bp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridisation to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridised to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (1) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central  
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly by detecting mutations or single nucleotide polymorphisms  
 CC (SNP's); and (1i) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation  
 CC status of many C residues to be determined simultaneously. ABQ13410-  
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention  
 CC  
 XX  
 SQ Sequence 874 BP; 144 A; 90 C; 320 G; 320 T; 0 U; 0 Other;  
 Query Match 47.7%; Score 62; DB 6; Length 874;  
 Best Local Similarity 71.8%; Pred. No. 2.6e-06;  
 Matches 94; Conservative 0; Mismatches 36; Indels 1; Gaps 1;  
 OY 1 TCCGCTTACTTCGGCTACCCCGGAAACCGCTTGCAGCTGTGGCGCTGATGGCGGATTCC 60  
 DB 341 TTCGTTATTTTCGGTTATTTTCGGGAATCGTTTCGAGTTGGTGGCGTTGATGGCGGATTTC 400  
 OY 61 GTGCTCTCCGACAGCCCGGCGCCACCT-GGGAGNAGTGTGAGCCCTGTGACCTTGCGCA 119  
 DB 401 GTGTTTTCGATGATTTTCGTTTATTTTGGGTGAGTGTGACGTTGTGATTTTCTGTA 460  
 OY 120 GGGACGCTGCT 130  
 DB 461 GGGACGTTGTT 471  
 RESULT 10  
 ABQ44405/c  
 ID ABQ44405 standard; DNA; 874 BP.  
 XX  
 AC ABQ44405;  
 XX

DT 12-JUL-2002 (first entry)  
 XX  
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30996.  
 XX  
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200218632-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PF 01-SEP-2001; 2001WO-EP010074.  
 XX  
 PR 01-SEP-2000; 2000DE-01043826.  
 PR 05-SEP-2000; 2000DE-01044543.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;  
 DR WPI; 2002-371829/40.  
 XX  
 PT Determining the degree of cytosine methylation in genomic DNA, useful for  
 PT diagnosis and prognosis, comprises selective hybridization of amplicons  
 PT from chemically treated DNA.  
 XX  
 PS Claim 12; 56bp + Sequence Listing; 56bp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridisation to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridised to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (1) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central  
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly by detecting mutations or single nucleotide polymorphisms  
 CC (SNP's); and (1i) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation  
 CC status of many C residues to be determined simultaneously. ABQ13410-  
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention  
 CC  
 XX  
 SQ Sequence 874 BP; 320 A; 320 C; 90 G; 144 T; 0 U; 0 Other;  
 Query Match 47.7%; Score 62; DB 6; Length 874;  
 Best Local Similarity 71.8%; Pred. No. 2.6e-06;  
 Matches 94; Conservative 0; Mismatches 36; Indels 1; Gaps 1;  
 OY 1 TCCGCTTACTTCGGCTACCCCGGAAACCGCTTGCAGCTGTGGCGCTGATGGCGGATTCC 60  
 DB 534 TTCGTTATTTTCGGTTATTTTCGGGAATCGTTTCGAGTTGGTGGCGTTGATGGCGGATTTC 475  
 OY 61 GTGCTCTCCGACAGCCCGGCGCCACCT-GGGAGNAGTGTGAGCCCTGTGACCTTGCGCA 119  
 DB 474 GTGTTTTCGATGATTTTCGTTTATTTTGGGTGAGTGTGACGTTGTGATTTTCTGTA 415  
 OY 120 GGGACGCTGCT 130  
 DB 414 GGGACGTTGTT 404  
 RESULT 11

AD34132  
 ID ADD34132 standard; DNA; 650 BP.  
 AC ADD34132;  
 DT 15-JAN-2004 (first entry)  
 DE Mouse mitochondrial DNA sequence SEQ ID NO:1910.  
 XX  
 KM de; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
 KM mitochondrial disease; oxidative phosphorylation dysfunction;  
 KM oxidative stress; apoptosis; aging.  
 XX  
 OS Mus musculus.  
 XX  
 PN W02003020220-A2.  
 XX  
 PD 13-MAR-2003.  
 XX  
 PF 30-AUG-2002; 2002WO-US027886.  
 XX  
 PR 30-AUG-2001; 2001US-0316323P.  
 XX  
 PR 31-AUG-2001; 2001CA-02356540.  
 XX  
 PA (UYEM-) UNIV EMORY.  
 PI Wallace DC, Levy S, Kerstam K, Procaccio V;  
 DR WPI; 2003-300821/29.  
 XX  
 PT Array containing probes for genes involved in mitochondrial biology,  
 PT useful for determining mitochondrial biology gene expression profiles for  
 PT use in diagnosing pathologies and identifying biochemical pathways.  
 XX  
 PS Claim 2; SEQ ID NO 1910; 201pp; English.  
 XX  
 CC The invention relates to a novel array comprising at least two isolated  
 CC nucleotide molecules, each molecule having a sequence capable of uniquely  
 CC hybridising to a nucleic acid molecule which is an expression product of  
 CC a gene involved in mitochondrial biology. The array comprises two or more  
 CC isolated nucleic acid molecules or spots, each molecule having a sequence  
 CC chosen from sequence of 994 human probes and 2046 mouse probes. An array  
 CC of the invention is useful for determining an expression profile of a  
 CC mouse or human sample containing nucleic acid, by contacting the array  
 CC with the sample under conditions allowing selective hybridisation, and  
 CC measuring hybridisation of nucleic acid in the sample to the array to  
 CC produce an expression profile. The array is also useful for determining  
 CC an expression profile of a first labelled sample containing nucleic acid  
 CC relative to a second, differently labelled sample containing nucleic  
 CC acid. The second sample is a reference or a standard. An array is useful  
 CC for determining an expression profile diagnostic of an energy-metabolism-  
 CC related physiological condition. An array of the invention is useful for  
 CC determining mitochondrial biology gene expression profiles of organisms,  
 CC such as human, mice and closely related species, tissue and organs of  
 CC such organisms, which are useful for determining expression profiles  
 CC diagnostic of energy metabolism-related physiological conditions,  
 CC diagnosing such physiological conditions, identifying biochemical  
 CC pathways, genes, and mutations involved in such physiological conditions,  
 CC identifying therapeutic agents useful for preventing and/or treating such  
 CC physiological conditions, evaluating and/or monitoring the efficacy of  
 CC such therapies, and creating and identifying animal models of human  
 CC energy metabolism-related physiological conditions. An array is also  
 CC useful for defining expression signatures or profiles for mitochondrial  
 CC diseases, as well as distinguishing clinical disorders that result from  
 CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,  
 CC apoptosis and aging. An array of the invention contains probes of genes  
 CC not previously recognised to participate in mitochondrial biology. The  
 CC sequences shown in ADD34132-ADD35260 represent murine mitochondrial DNA  
 CC clones used to make the probes of the invention. Some sequences are not  
 CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,  
 CC 1906, 2408 and 2643.  
 XX  
 SQ Sequence 650 BP; 143 A; 172 C; 170 G; 165 T; 0 U; 0 Other;

Query Match 39.2%; Score 51; DB 10; Length 650;  
 Best Local Similarity 67.2%; Pred. No. 0.0014;  
 Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;  
 QY 1 TCCGCTACTCTCGGCTACCCCGGAAACGCTTCGAGCTGTGGCGCTGATGGCGATTCC 60  
 DB 173 TCTCTCTTCTGCGAAGCCGGGCAATGCTGTGAGCTGTGAAACAGATGGCAGATTAAG 232  
 QY 61 GTGCTCTCCGACAGCCCCCGGCACTTGG-GAGNAGTGTGAGGCTGTGACTTGGCA 119  
 DB 233 TTGCTCTCCAAAGACCAAGACTTCACTGAGCGCAACTGTGATGCTCTCGCTCGCG 292  
 QY 120 GGAGCGCT 127  
 DB 293 GGAGCGCT 300  
 RESULT 12  
 AAT33800/C  
 ID AAT33800 standard; cDNA to mRNA; 3288 BP.  
 AC AAT33800;  
 XX  
 DT 05-DEC-1996 (first entry)  
 XX  
 DE Murine interleukin-17 receptor cDNA.  
 XX  
 KM Interleukin-17 receptor; IL-17R; autoimmune disease; allergy; asthma;  
 KM graft rejection; inflammation; therapy; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 121..2715  
 FT /\*tag= a  
 FT sig\_peptide 121..213  
 FT /\*tag= b  
 FT mat\_peptide 214..2712  
 FT /\*tag= c  
 XX  
 PN W09629408-A1.  
 XX  
 PD 26-SEP-1996.  
 XX  
 PF 21-MAR-1996; 96WO-US004018.  
 XX  
 PR 23-MAR-1995; 95US-00410535.  
 XX  
 PR 07-AUG-1995; 95US-00538765.  
 XX  
 PA (IMMV ) IMMUNEX CORP.  
 XX  
 PI Yao Z, Spriggs MK, Fanslow WC;  
 DR WPI; 1996-443184/44.  
 DR P-PSDB; AAM04184.  
 XX  
 PT DNA encoding interleukin-17 receptor - useful for regulating immune and  
 PT inflammatory responses, or to suppress graft rejection.  
 XX  
 PS Claim 1; Page 25-29; 52pp; English.  
 XX  
 CC A cDNA clone (AAT33800) codes for murine interleukin-17 receptor (IL-17R)  
 CC (AAM04184), a type I transmembrane protein. A cDNA library of murine  
 CC thymoma B14 cells was used to transform CV1/EBNA cells and the prods.  
 CC were screened for binding to a fusion between the FC region of human IgG1  
 CC (see also AAM02305) and the viral IL-17 homologue HVS13 (AAM02387). A  
 CC positive clone contained a 3.2 kb insert that included the IL-17R open  
 CC reading frame. The cDNA can be used to produce (pref. soluble)  
 CC recombinant IL-17R in transformed host cells or to transfect a tissue or  
 CC organ as a means of suppressing graft rejection. It was also used to  
 CC identify a cDNA clone (AAT33801) coding for human IL-17R (AAM04185)



CC polypeptides have immunoregulatory activity. They can be used for  
CC inhibiting T cell proliferation, or for inhibiting T cell activation. In  
CC particular they can be used for preventing or treating organ or graft  
CC rejection, autoimmune disease, allergy or asthma. They can also be used  
CC for the prevention or treatment of inflammatory disease in which  
CC activated T cells play a role or for inhibiting B cell proliferation or  
CC immunoglobulin secretion. The IL-17 polypeptides can also be used as  
CC immunogens, reagents in in vitro assays, or as binding agents for  
CC affinity purification procedures

XX SQ Sequence 3288 BP; 733 A; 928 C; 906 G; 721 T; 0 U; 0 Other;

Query Match 28.3%; Score 36.8; DB 2; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 5.4; Mismatches 0; Gaps 0;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 TCCGCTTACTCGGCTACCCCGGAAACCGCTTCGAGCTGTGCGCTGATGCGGATTCC 60  
DB 2552 TCTCTCTCCGAGGACACTCGGGGGGCTGCGGGAGCTCTGAGATGTAGCCCTGTGCC 2493  
QY 61 GTGCTTCCGACAGCCCGGCCCCACCTGGAGAGNAGTGTACCGCTCGTACCTTCCGAG 120  
DB 2492 GACTGCACCGACTCGCGCTCTCTCTCTCAAGAGGTGTGACGCCCTCGAGAGACCACTCT 2433  
QY 121 GGACGCTGC 129  
DB 2432 GGCTTCGGC 2424

RESULT 15  
AA51987/c  
ID AA51987 standard; cDNA to mRNA; 3288 BP.  
AC AA51987;  
XX 04-DEC-2000 (first entry)  
DT  
XX Murine interleukin-17 receptor coding sequence.  
DE  
XX IL-17R; interleukin-17 receptor; soluble; CTLA-8; Herpesvirus saimiri;  
KM HSV13; graft rejection; suppressor; immunosuppressive; anti-allergic;  
KM anti-asthmatic; ss.  
XX  
OS Mus sp.  
XX  
XX Key Location/Qualifiers  
FH 121..2715  
FT /\*tag= a  
FT /product= "IL-17\_receptor"  
FT sig\_peptide 121..213  
FT /\*tag= b  
FT mat\_peptide 214..2712  
FT /\*tag= c  
XX  
XX US6100235-A.  
XX  
XX 08-AUG-2000.  
PD  
XX  
XX 11-FEB-1998; 98US-00022260.  
XX  
XX 23-MAR-1995; 95US-00410535.  
PR 07-AUG-1985; 95US-00538765.  
PR 21-MAR-1996; 96US-00620694.  
XX  
XX (IMMV ) IMMUNEX CORP.  
XX  
XX Fanslow WC, Spriggs MK, Yao Z;  
PI WPI; 2000-548298/50.  
DR P-PSDB; AAY97130.  
XX  
XX Regulating, treating or preventing immune or inflammatory response in a  
PT mammal, especially organ or graft rejection, allergy or asthma, comprises

PT administering interleukin-17 receptors.  
XX  
XX Example 2; Col 19-26; 26pp; English.  
PS  
XX A novel interleukin-17 receptor (IL-17R) was identified by screening a  
CC cDNA library from T cell thymoma EL4 cells, which were identified as  
CC binding fusion proteins comprising human immunoglobulin G1 (IgG1) Fc  
CC region and soluble IL-17 (CTLA-8) protein or a homologous Herpesvirus  
CC saimiri (HSV) protein, designated HSV13. Regulating an immune or  
CC inflammatory response in a mammal comprises administering soluble IL-17R.  
CC Soluble IL-17R fragments comprise residues 1-322 of the murine IL-17R,  
CC residues 1-320 of the human IL-17R, and fragments of the extracellular  
CC domain that bind IL-17. The method is useful for regulating an  
CC immunoreponse, for treating or preventing diseases like allergy, asthma  
CC and autoimmune diseases, and for suppressing rejection of grafted organs  
CC or tissues in the recipient

XX SQ Sequence 3288 BP; 733 A; 928 C; 906 G; 721 T; 0 U; 0 Other;

Query Match 28.3%; Score 36.8; DB 3; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 5.4; Mismatches 0; Gaps 0;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 TCCGCTTACTCGGCTACCCCGGAAACCGCTTCGAGCTGTGCGCTGATGCGGATTCC 60  
DB 2552 TCTCTCTCCGAGGACACTCGGGGGGCTGCGGGAGCTCTGAGATGTAGCCCTGTGCC 2493  
QY 61 GTGCTTCCGACAGCCCGGCCCCACCTGGAGAGNAGTGTACCGCTCGTACCTTCCGAG 120  
DB 2492 GACTGCACCGACTCGCGCTCTCTCTCTCAAGAGGTGTGACGCCCTCGAGAGACCACTCT 2433  
QY 121 GGACGCTGC 129  
DB 2432 GGCTTCGGC 2424

Search completed: June 7, 2005, 22:14:36  
Job time : 130.188 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 06:21:30 ; Search time 41.4159 Seconds  
(without alignments)  
5136.091 Million cell updates/sec

Title: US-10-071-174A-37

Perfect score: 130

Sequence: 1 tcgcgcaccctcgcgcacc.....acctcgcagcgacgcgcgc 130

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents, NA.\*

2: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.6	28.9	1704	US-09-902-540-4478	Sequence 4478, Ap
C 2	37.6	28.9	28172	US-09-902-540-1221	Sequence 1221, Ap
C 3	36.8	28.3	3288	US-08-620-694A-1	Sequence 1, Appli
C 4	36.8	28.3	3288	US-09-022-255-1	Sequence 1, Appli
C 5	36.8	28.3	3288	US-08-978-773-1	Sequence 1, Appli
C 6	36.8	28.3	3288	US-09-022-253-1	Sequence 1, Appli
C 7	36.8	28.3	3288	US-09-022-253-1	Sequence 1, Appli
C 8	36.8	28.3	3288	US-09-022-253-1	Sequence 1, Appli
C 9	36.8	28.3	3288	US-09-022-253-1	Sequence 1, Appli
C 10	36.8	28.3	3288	US-09-022-253-1	Sequence 1, Appli
C 11	36.8	28.3	3288	US-09-549-679-1	Sequence 1, Appli
C 12	36.8	28.3	3288	US-09-902-540-6511	Sequence 6511, Ap
C 13	36.8	28.3	3288	US-09-902-540-472	Sequence 472, App
C 14	36.8	28.3	3288	US-09-270-767-13105	Sequence 13105, A
C 15	36.8	28.3	3288	US-08-858-003-29	Sequence 29, Appli
C 16	36.8	28.3	3288	US-09-078-166-29	Sequence 29, Appli
C 17	36.8	28.3	3288	US-08-997-467-29	Sequence 29, Appli
C 18	36.8	28.3	3288	US-09-489-039A-6329	Sequence 6329, Ap
C 19	36.8	28.3	3288	US-09-489-039A-6050	Sequence 6050, Ap
C 20	36.8	28.3	3288	US-08-804-227C-7	Sequence 7, Appli
C 21	36.8	28.3	3288	US-08-804-198-1	Sequence 1, Appli
C 22	36.8	28.3	3288	US-09-949-016-15448	Sequence 15448, A
C 23	36.8	28.3	3288	US-09-724-797-133	Sequence 13, Appli
C 24	36.8	28.3	3288	US-09-902-540-8389	Sequence 8389, Ap
C 25	36.8	28.3	3288	US-08-713-928B-8	Sequence 8, Appli
C 26	36.8	28.3	3288	US-09-191-171-4	Sequence 4, Appli
C 27	36.8	28.3	3288	US-09-385-707-4	Sequence 4, Appli

28	31.4	24.2	2155	4	US-09-639-696C-4	Sequence 4, Appli
29	31.4	24.2	2155	4	US-09-917-254-30	Sequence 30, Appli
30	31.4	24.2	4480	3	US-09-191-171-7	Sequence 7, Appli
31	31.4	24.2	4480	3	US-09-385-707-7	Sequence 7, Appli
32	31.4	24.2	6200	3	US-09-439-923-1	Sequence 1, Appli
33	31.4	24.2	6200	3	US-09-711-202A-1	Sequence 1, Appli
34	31.4	24.2	6200	3	US-09-711-202A-1	Sequence 1, Appli
35	31.4	24.2	6238	4	US-09-639-696C-6	Sequence 6, Appli
36	31.4	24.2	7186	4	US-09-513-999C-688	Sequence 688, App
37	30.8	23.7	753	4	US-09-902-540-4067	Sequence 4067, Ap
38	30.8	23.7	870	4	US-09-252-951A-1268	Sequence 1268, Ap
39	30.8	23.7	5481	4	US-09-902-540-7351	Sequence 7351, Ap
40	30.8	23.7	6402	4	US-09-902-540-7351	Sequence 702, App
41	30.8	23.7	21511	4	US-09-902-540-1201	Sequence 1201, Ap
42	30.8	23.7	897	3	US-09-434-288-6	Sequence 6, Appli
43	30.6	23.5	1320	4	US-09-727-238-1	Sequence 1, Appli
44	30.6	23.5	1365	3	US-09-319-892-1	Sequence 1, Appli
45	30.6	23.5	1365	3	US-09-319-892-1	Sequence 1, Appli

## ALIGNMENTS

```
RESULT 1
US-09-902-540-4478/C
Sequence 4478, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(115849)B
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 4478
LENGTH: 1704
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-4478

Query Match      28.9%; Score 37.6; DB 4; Length 1704;
Best Local Similarity 56.0%; Pred. No. 0.3;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Cy 5 CTTACCTCGGCTACCCGCGGAAAGCGCTTCGAGCGTGTGGCGGCTGATGCGGATTCGCGC 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 CCGGCTTCACCCACAGCGGCGCGCGCTTCGCGGCGCGCGCGCGCGCTTCGCGGCGCGC 821
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cy 65 TCTCCGACAGCCCGCGGCGCGCGCTTCGAGGAGGAGTGTGACGCTTCGAGCGGAGC 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 CCACCTTCATGACGAGCGCGCGCGCGCGCGGAGATGCTCCACAGCGGCGCTCGAGAAC 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cy 125 GCTGC 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 GCTGC 756

RESULT 2
US-09-902-540-1221
Sequence 1221, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(115849)B
```

;; CURRENT APPLICATION NUMBER: US/09/902,540  
;; CURRENT FILING DATE: 2001-07-10  
;; PRIOR APPLICATION NUMBER: 60/217,883  
;; PRIOR FILING DATE: 2000-07-10  
;; NUMBER OF SEQ ID NOS: 16825  
;; SEQ ID NO 1221  
;; LENGTH: 28172  
;; TYPE: DNA  
;; ORGANISM: Myxococcus xanthus  
US-09-902-540-1221

Query Match 28.9%; Score 37.6; DB 4; Length 28172;  
Best Local Similarity 56.0%; Pred. No. 0.43;  
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 5 CCTACCTGGCTACCCCGGGAACCGCTTCGAGCTGTGCGCTGATGGCGGATTCCGTC 64  
Db 18175 CCGGTTACCCACGCGGACCGCTTCGAGCGCGCGCGCGCTCGGCTACTCCGCGC 18234  
Qy 65 TCTCCGACAGCCCGCCGACCTTGAGAGNAGTGTGACGCTCTGACCTTCGACGGGAC 124  
Db 18235 CCAGTTTATGAGCGCGCGCGCGCGAGATGCTCCACACGCGCGCTTGAGAGAAC 18294  
Qy 125 GCTGC 129  
Db 18295 GCTGC 18299

## RESULT 3

US-08-620-694A-1/c  
; Sequence 1, Application US/08620694A  
; Patent No. 5869286

## GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,694A  
FILING DATE: 21 MARCH 1996  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206)

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3288 base pairs  
TYPE: nucleic acid

;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Mouse  
;; STRAIN: HVS13 receptor  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 121..2715  
US-08-620-694A-1

Query Match 28.3%; Score 36.8; DB 2; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 0.54;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 TCCGCTACTCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGAGCGCTGATGGCGGATTCC 60  
Db 2552 TCTTCTCTCCGTGACCACTCGGGGCGCTGCGGGAGCTCTCGAGATGATACCTGTGTC 2493  
Qy 61 GTGCTTCCGACAGCCCGCGCCGACCTTGAGAGNAGTGTGAGCGCTGTGACCTTTCGAG 120  
Db 2492 GACTGACCGAGCTCCGCTGCTCTCTCAGAGGCTGTGACGCCCTTGAGGACCACTCT 2433  
Qy 121 GAGCGTGC 129  
Db 2432 GCGCTCGGC 2424

## RESULT 4

US-09-022-255-1/c  
; Sequence 1, Application US/09022255  
; Patent No. 6072033

## GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,255  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/620,694  
FILING DATE: 21 MARCH 1996  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995

## ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206)

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3288 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse  
STRAIN: HVS13 receptor  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 121..2715  
US-09-022-255-1

Query Match 28.3%; Score 36.8; DB 3; Length 3288;

Best Local Similarity 55.0%; Pred. No. 0.54; Mismatches 58; Indels 0; Gaps 0;

Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 TCCGCTACCTCGGCTACCCCGGAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 60  
Db 2552 TCTCTCTCCGAGGCACTCGGGGGGCTGCGGGAGCTCTGAGATGTAGCCCTGTGTC 2493  
Qy 61 GTGCTCTCCGACAGCCCGGCGCCACCTGGAGAGNAGTGATGACGCTGTGACCTTCGAG 120  
Db 2492 GACTGCACCGACTGCGCTGCTCTCTCTCAAGAGGTGTGACAGCCCTCGAGACCACTCT 2433  
Qy 121 GGACGCTGC 129  
Db 2432 GGCTCTGGC 2424

## RESULT 5

US-09-022-696-1/c

Sequence 1, Application US/09022696

Patent No. 6072037

GENERAL INFORMATION:

APPLICANT: Yao, Zhenbin

APPLICANT: Farnlow, William

TITLE OF INVENTION: No. 6072037e1 Receptor That Binds IL-17

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Power Macintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for Apple, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/022.696

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/620, 694

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/410, 535

FILING DATE: 23 MARCH 1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,695

REFERENCE/DOCKET NUMBER: 2617-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3288 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse  
STRAIN: HVS13 receptor  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 121..2715  
US-09-022-696-1

Query Match 28.3%; Score 36.8; DB 3; Length 3288;

Best Local Similarity 55.0%; Pred. No. 0.54; Mismatches 58; Indels 0; Gaps 0;

Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 TCCGCTACCTCGGCTACCCCGGAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 60  
Db 2552 TCTCTCTCCGAGGCACTCGGGGGGCTGCGGGAGCTCTGAGATGTAGCCCTGTGTC 2493  
Qy 61 GTGCTCTCCGACAGCCCGGCGCCACCTGGAGAGNAGTGATGACGCTGTGACCTTCGAG 120  
Db 2492 GACTGCACCGACTGCGCTGCTCTCTCTCAAGAGGTGTGACAGCCCTCGAGACCACTCT 2433  
Qy 121 GGACGCTGC 129  
Db 2432 GGCTCTGGC 2424

## RESULT 6

US-08-978-773-1/c

Sequence 1, Application US/08978773

Patent No. 6083906

GENERAL INFORMATION:

APPLICANT: Trout, Anthony

TITLE OF INVENTION: Method of Regulating Nitric Oxide Production

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple PowerMacintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978, 773

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 60/052, 525

FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: 2623-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3288 base pairs

TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse  
IMMEDIATE SOURCE:  
CLONE: IL-17 receptor  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 121..2712  
US-08-978-773-1

Query Match 28.3%; Score 36.8; DB 3; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 0.54;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 TCCGCTTACCTCGGCTACCCCGGAAACCGCTTTCAGAGTGTGGCGCTGATGCGGATTC 60  
Db 2552 TCTCTCTCCGTAGACCACTCGGGGGGGCTGGGGAGCTCTTGAGATGAGCCCTGTCC 2493  
Qy 61 GTGCTTCCGACAGCCCGGCCCCCACCCTGGAGAGNAGTGATGACGCTCGTGAACCTTCGAG 120  
Db 2492 GACTGCACCGACTGCGCGCTGCTCTCTTCAGAGGGGTGTGACGCCCTCGAGGACCACTCT 2433  
Qy 121 GGACGCTGC 129  
Db 2432 GGCTTCGGC 2424

## RESULT 7

US-09-022-253-1/c  
Sequence 1, Application US/09022253  
Patent No. 6096305

GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022.253

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/620.694  
FILING DATE: 21-MARCH-1996  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995

ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430

TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3288 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse  
STRAIN: HVS13 receptor  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 121..2715  
US-09-022-253-1

Query Match 28.3%; Score 36.8; DB 3; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 0.54;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 TCCGCTTACCTCGGCTACCCCGGAAACCGCTTTCAGAGTGTGGCGCTGATGCGGATTC 60  
Db 2552 TCTCTCTCCGTAGACCACTCGGGGGGGCTGGGGAGCTCTTGAGATGAGCCCTGTCC 2493  
Qy 61 GTGCTTCCGACAGCCCGGCCCCCACCCTGGAGAGNAGTGATGACGCTCGTGAACCTTCGAG 120  
Db 2492 GACTGCACCGACTGCGCGCTGCTCTCTTCAGAGGGGTGTGACGCCCTCGAGGACCACTCT 2433  
Qy 121 GGACGCTGC 129  
Db 2432 GGCTTCGGC 2424

## RESULT 8

US-09-022-260-1/c  
Sequence 1, Application US/09022260  
Patent No. 6100235

GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022.260

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)587-0430  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3288 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse  
STRAIN: HVS13 receptor  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 121..2715  
US-09-022-260-1

Query Match 28.3%; Score 36.8; DB 3; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 0.54;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 TCCGCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 60  
Db 2552 TCTCTCTCCGAGAGCACTCGGGGGGGCTGCGGGAGCTCTCGAATGTAGCCCTGGTCC 2493  
Qy 61 GTGCTCTCGAAGCGCCCGCCCACTGGAGAGAGTGATGAGCTGTGACCTTCGAG 120  
Db 2492 GACTGCACCGACTGCGCTGCTCTCTCAAGAGGTGTGACAGCCTCGAGAGCACCTCT 2433  
Qy 121 GGACGCTGC 129  
Db 2432 GGCTCTGGC 2424

RESULT 9  
US-09-022-259-1/c  
Sequence 1, Application US/09022259  
Patent No. 6191104  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanelow, William  
TITLE OF INVENTION: No. 6191104e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,259  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)587-0430  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3288 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse  
STRAIN: HVS13 receptor  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 121..2715  
US-09-022-259-1

Query Match 28.3%; Score 36.8; DB 3; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 0.54;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 TCCGCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGGCGATTCC 60  
Db 2552 TCTCTCTCCGAGAGCACTCGGGGGGGCTGCGGGAGCTCTCGAATGTAGCCCTGGTCC 2493  
Qy 61 GTGCTCTCGAAGCGCCCGCCCACTGGAGAGAGTGATGAGCTGTGACCTTCGAG 120  
Db 2492 GACTGCACCGACTGCGCTGCTCTCTCAAGAGGTGTGACAGCCTCGAGAGCACCTCT 2433  
Qy 121 GGACGCTGC 129  
Db 2432 GGCTCTGGC 2424

RESULT 10  
US-09-022-257-1/c  
Sequence 1, Application US/09022257  
Patent No. 6197525  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanelow, William  
TITLE OF INVENTION: No. 6197525e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,257  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:

FILING DATE: 23 MARCH  
ATTORNEY/AGENT INFORMATION:

[illegible]

GAGCCTGGCCGCGTGGTGAAGGACGT 121/



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 06:21:09 ; Search time 122.203 Seconds  
(without alignments)  
6090.078 Million cell updates/sec

Title: US-10-071-174A-37  
Perfect score: 130  
Sequence: 1 tccgcctactctgcgtacc...acctcgcaggagcctgcct 130

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4285876 seqs, 2862396386 residues

Total number of hits satisfying chosen parameters: 8571752

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/prodata/1/pubpna/PCF\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 5: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 7: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 8: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 12: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 13: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 14: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 15: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 16: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 17: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 18: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 19: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 20: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 21: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 22: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 23: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 24: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 25: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	* Query Match	Length	ID	Description
1	114.8	88.3	887	16	US-10-071-174-1
2	100.8	77.5	548	9	US-09-764-847-311
3	100.8	77.5	548	14	US-10-092-154-311
4	86.8	66.8	522	9	US-09-912-589-3
5	73.2	56.3	874	20	US-10-363-345A-30993
6	73.2	56.3	874	20	US-10-363-345A-30994
7	62	47.7	874	20	US-10-363-345A-30995
8	62	47.7	874	20	US-10-363-345A-30996

C	9	36.8	28.3	3288	19	US-10-742-161-1	Sequence 1, Appl1
C	10	36.8	28.3	3288	19	US-10-742-172-1	Sequence 1, Appl1
C	11	36.4	28.0	3630	19	US-10-437-963-12593	Sequence 12593, A
C	12	36.2	27.8	632	20	US-10-425-115-169387	Sequence 169387, A
C	13	36	27.7	486	15	US-10-156-761-1614	Sequence 1614, Ap
C	14	36	27.7	9025608	15	US-10-156-761-1614	Sequence 1, Appl1
C	15	35.4	27.2	2412	17	US-10-260-238-1034	Sequence 1034, Ap
C	16	35.4	27.2	2959	19	US-10-437-963-62213	Sequence 62213, A
C	17	34.8	26.8	1783	19	US-10-437-963-54971	Sequence 54971, A
C	18	34.6	26.6	1095	19	US-10-437-963-93303	Sequence 93303, A
C	19	34	26.2	1595	18	US-10-424-599-103250	Sequence 103250, A
C	20	33.6	25.8	1077	17	US-10-260-238-313	Sequence 313, App
C	21	33.4	25.7	771	19	US-10-767-701-1406	Sequence 1406, Ap
C	22	33.4	25.7	878	19	US-10-437-963-65205	Sequence 65205, A
C	23	33	25.4	852	20	US-10-425-115-88641	Sequence 88641, A
C	24	33	25.4	1911	14	US-10-152-886-30	Sequence 88641, A
C	25	32.6	25.1	882	17	US-10-402-842-34	Sequence 30, Appl1
C	26	32.6	25.1	47988	17	US-10-402-842-1	Sequence 34, Appl1
C	27	32.4	24.9	385	19	US-10-767-701-28894	Sequence 1, Appl1
C	28	32.4	24.9	1194	15	US-10-156-761-3451	Sequence 28894, A
C	29	32.2	24.8	539	19	US-10-437-963-82849	Sequence 3451, Ap
C	30	32.2	24.8	1866	19	US-10-437-963-84852	Sequence 82849, A
C	31	32.2	24.8	17083	19	US-10-475-970-4	Sequence 84852, A
C	32	32	24.6	792	19	US-10-437-963-2791	Sequence 4, Appl1
C	33	32	24.6	877	20	US-10-425-115-171782	Sequence 2791, Ap
C	34	32	24.6	1489	19	US-10-437-963-22814	Sequence 171782, A
C	35	32	24.6	1786	20	US-10-739-930-5058	Sequence 22814, A
C	36	32	24.6	1806	20	US-10-739-930-5055	Sequence 5058, Ap
C	37	32	24.6	3807	17	US-10-282-122A-26053	Sequence 5055, Ap
C	38	31.8	24.5	993	19	US-10-437-963-10729	Sequence 26053, A
C	39	31.8	24.5	1010	10	US-09-735-056-29	Sequence 10729, A
C	40	31.8	24.5	2107	20	US-10-425-115-62599	Sequence 29, Appl1
C	41	31.6	24.3	560	18	US-10-424-599-55478	Sequence 62599, A
C	42	31.6	24.3	2079	15	US-10-156-761-1755	Sequence 55478, A
C	43	31.6	24.3	16018	9	US-09-764-869-1384	Sequence 1755, Ap
C	44	31.6	24.3	16018	14	US-10-091-504-1384	Sequence 1384, Ap
C	45	31.6	24.3	16018	17	US-10-227-577-1384	Sequence 1384, Ap

#### ALIGNMENTS

RESULT 1  
US-10-071-174-1  
Sequence 1, Application US/10071174  
Publication No. US2003017671A1  
GENERAL INFORMATION:  
APPLICANT: REED, JOHN C.  
APPLICANT: KE, NING  
APPLICANT: GODZIK, ADAM  
TITLE OF INVENTION: APOPTOSIS MODULATOR BCL-B AND METHODS FOR MAKING AND  
FILE REFERENCE: 087102-0272558  
CURRENT APPLICATION NUMBER: US/10/071, 174  
CURRENT FILING DATE: 2002-02-07  
PRIORITY APPLICATION NUMBER: 60/267,166  
PRIORITY FILING DATE: 2001-02-07  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 887  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-071-174-1

Query Match: 88.3%, Score 114.8, DB 16, Length 887;  
Best Local Similarity 96.9%; Pred. No. 1.9e-25;  
Matches 127; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
OY 1 TCCGCTACTCTGCGCTACCCCGGAGACCGCTTGAGCTGGTGGCGCTGATGGCGGATTC 60  
DB 242 TCCGCTACTCTGCGCTACCCCGGAGACCGCTTGAGCTGGTGGCGCTGATGGCGGATTC 301



```
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363.345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 30993
; LENGTH: 874
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 30993
US-10-363-345A-30993
```

```
Query Match          56.3%; Score 73.2; DB 20; Length 874;
Best Local Similarity 77.1%; Pred. No. 7.7e-13;
Matches 101; Conservative 0; Mismatches 29; Indels 1; Gaps 1;
```

```
Qy 1 TCCGCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGGCGGATTCC 60
    |||
Db 534 TCCGCTACCTCGGCTACCTACCCGGAACCGCTTCGAGCTGTGGCGCTGATGGCGGATTCC 475
Qy 61 GTGCTTCGACAGCCCCCGGCCCACTTGGA-GNAGTGTGACGCTGTGACCTTCCGA 119
    |||
Db 474 GTACTCTCCGACACCCCGACCCCACTTAACAAATAATATAGCTGTGATACCTTCCGA 415
Qy 120 GGGACGCTGCT 130
    |||
Db 414 AAAACGCTACT 404
```

```
RESULT 6
US-10-363-345A-30994
; Sequence 30994, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363.345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 30994
; LENGTH: 874
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 30994
US-10-363-345A-30994
```

```
Query Match          56.3%; Score 73.2; DB 20; Length 874;
Best Local Similarity 77.1%; Pred. No. 7.7e-13;
Matches 101; Conservative 0; Mismatches 29; Indels 1; Gaps 1;
```

```
Qy 1 TCCGCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGGCGGATTCC 60
    |||
Db 341 TCCGCTACCTCGGCTACCTACCCGGAACCGCTTCGAGCTGTGATAGCTAATACGATTCC 400
Qy 61 GTGCTTCGACAGCCCCCGGCCCACTTGGA-GNAGTGTGACGCTGTGACCTTCCGA 119
    |||
Db 401 GTACTCTCCGACACCCCGACCCCACTTAACAAATAATATAGCTGTGATACCTTCCGA 460
Qy 120 GGGACGCTGCT 130
    |||
```

```
Db 461 AAAACGCTACT 471
```

```
RESULT 7
US-10-363-345A-30995
; Sequence 30995, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363.345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 30995
; LENGTH: 874
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 30995
US-10-363-345A-30995
```

```
Query Match          47.7%; Score 62; DB 20; Length 874;
Best Local Similarity 71.8%; Pred. No. 1.9e-09;
Matches 94; Conservative 0; Mismatches 36; Indels 1; Gaps 1;
```

```
Qy 1 TCCGCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGGCGGATTCC 60
    |||
Db 341 TTGCTTATTTTCGTTATTTTCGGAATCGTTTCAGATGGTGGGCTGATGCGGATTTC 400
Qy 61 GTGCTTCGACAGCCCCCGGCCCACTTGGA-GGAGNAGTGTGACGCTGTGACCTTCCGA 119
    |||
Db 401 GTGTTTTCGATGATTTTCGTTTATTTTGGGTAGAGTGTGACGCTGTGATTTCCGA 460
Qy 120 GGGACGCTGCT 130
    |||
Db 461 GGGACGTTGTT 471
```

```
RESULT 8
US-10-363-345A-30996/c
; Sequence 30996, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363.345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 30996
; LENGTH: 874
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 30996
US-10-363-345A-30996
```

```
Query Match          47.7%; Score 62; DB 20; Length 874;
```

```
Best Local Similarity 71.8%; Pred. No. 1.9e-09;
```

```
Matches 94; Conservative 0; Mismatches 36; Indels 1; Gaps 1;
Qy 1 TCCGCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGGCGCTGATGGCGGATTCC 60
    |||
Db 534 TTGCTTATTTTCGTTATTTTCGGAATCGTTTCAGATGGTGGGCTGATGCGGATTTC 475
```

Qy 61 GTGCTCTCCGACAGCCCGGCCCCACT-GGAGAGNAGTGTGACCGCTGTGACCTTCCGA 119  
Db 474 GTGCTTTCGATGATTTTCGGTTTATTTGGGTAGAGTGTACCTTGTGATTTTCGTA 415  
Qy 120 GGGAGCGTGCT 130  
Db 414 GGGACGTTGTT 404

## RESULT 9

US-10-742-161-1/c  
; Sequence 1, Application US/10742161  
; Publication No. US20040120898A1  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; Spriggs, Melanie  
; Fanslow, William  
; TITLE OF INVENTION: Novel Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/742,161  
; FILING DATE: 18-Dec-2003  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/620,694  
; FILING DATE: 21 MARCH 1996  
; APPLICATION NUMBER: USSN 08/538,765  
; FILING DATE: 7 AUGUST 1995  
; APPLICATION NUMBER: USSN 08/410,535  
; FILING DATE: 23 MARCH 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,695  
; REFERENCE/DOCKET NUMBER: 2617-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206)  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3288 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Mouse  
; STRAIN: HVS13 receptor  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 121..2715  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-742-161-1

Query Match 28.3%; Score 36.8; DB 19; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 0.073; Mismatches 58; Indels 0; Gaps 0;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
Qy 1 TCCGCTACTCTGCTACCCGGGAAACGCTTGCAGTGTGTGGCCGTGATGCGGATTCC 60

Db 2552 TCCCTCTCCGAGACCACTCGGGGGGCTGCGGGAGCTCTCGAGATGTACCCCTGATCC 2493  
Qy 61 GTGCTCTCCGACAGCCCGGCGCCACCTGGGAGAGNAGTGTAGCGCTGTGACCTTCCGAG 120  
Db 2492 GACTGACAGCACTGCGCTGCTCTCTCAAGAGGTGTGACGCCCTGAGGACCACTCT 2433  
Qy 121 GGACGCTGC 129  
Db 2432 GGCTCTGGC 2424

## RESULT 10

US-10-742-372-1/c  
; Sequence 1, Application US/10742372  
; Publication No. US20040120899A1  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; Spriggs, Melanie  
; Fanslow, William  
; TITLE OF INVENTION: Novel Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/742,372  
; FILING DATE: 18-Dec-2003  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/620,694  
; FILING DATE: 21 MARCH 1996  
; APPLICATION NUMBER: USSN 08/538,765  
; FILING DATE: 7 AUGUST 1995  
; APPLICATION NUMBER: USSN 08/410,535  
; FILING DATE: 23 MARCH 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,695  
; REFERENCE/DOCKET NUMBER: 2617-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206)  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3288 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Mouse  
; STRAIN: HVS13 receptor  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 121..2715  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-742-372-1

Query Match 28.3%; Score 36.8; DB 19; Length 3288;  
Best Local Similarity 55.0%; Pred. No. 0.073; Mismatches 58; Indels 0; Gaps 0;  
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 TCCGCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGTGGCCGCTGATGCGGATTC 60  
DB 2552 TCTCTCTCCGTGAGCCATCCGGGGGGCTGCGGGAGCTCTCGAGATGTAGCCCTCGTCC 2493  
QY 61 GTGCTTCGACAGCCCGGCGCCCACTGTGGAGNAGTGTGACGCTGTGACCTTCGAG 120  
DB 2492 GACTGACACGACCTGCGCTGTCTCTCTCAGAGGTGTGACGCTTCGAGACCACTCT 2433  
QY 121 GGACGCTGC 129  
DB 2432 GGCTCGGC 2424

RESULT 11  
US-10-437-963-12593/c  
; Sequence 12593, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Mu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 12593  
; LENGTH: 3630  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_18706C.1  
US-10-437-963-12593

Query Match 28.0%; Score 36.4; DB 19; Length 3630;  
Best Local Similarity 55.1%; Pred. No. 0.095; Indels 0; Gaps 0;  
Matches 70; Conservative 0; Mismatches 57;

QY 3 CGCTACCTCGGCTACCCCGGAAACCGCTTCGAGCTGTGTGGCCGCTGATGCGGATTCGT 62  
DB 972 CGCAGAGCTCTCTGCGCGGACGAGGTTCCCGCGCGCGCTGCTCTCTCTC 913  
QY 63 GCTCTCCGACAGCCCGGCGCCCACTGTGGAGNAGTGTGACGCTGTGACCTTCGAGG 122  
DB 912 GCCGACCGCGCGCGCGGACGAGCTGTGCGGTAACCGCTTCGCGCGCGCGGAG 853  
QY 123 ACGCTGC 129  
DB 852 TCGCGGC 846

RESULT 12  
US-10-425-115-169387  
; Sequence 169387, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 169387  
; LENGTH: 632  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_86067C.1  
US-10-425-115-169387

Query Match 27.8%; Score 36.2; DB 20; Length 632;  
Best Local Similarity 55.7%; Pred. No. 0.13; Indels 0; Gaps 0;  
Matches 68; Conservative 0; Mismatches 54;

QY 9 CCGCTACCTACCCCGGAAACCGCTTCGAGCTGTGTGGCCGCTGATGCGGATTCGCTCTC 68  
DB 391 CCGCGCGCGGACAGCGCGGCGCACCTCGCTCTCGCGGCGCTCGCGCGCGCGCTCGC 450  
QY 69 CGACAGCCCGGCGCCCACTGTGGAGNAGTGTGACGCTGTGACCTTCGAGGAGCGTG 128  
DB 451 CGCGCGCGCGCGCGCGCGCGCTGTGCGGAGAGCGCTGCGCGCTGCGCGGCTCGCG 510  
QY 129 CT 130  
DB 511 CT 512

RESULT 13  
US-10-156-761-1614  
; Sequence 1614, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1614  
; LENGTH: 486  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1) ..(486)  
US-10-156-761-1614

Query Match 27.7%; Score 36; DB 15; Length 486;  
Best Local Similarity 59.4%; Pred. No. 0.15; Indels 0; Gaps 0;  
Matches 60; Conservative 0; Mismatches 41;

QY 12 CGGCTACCCCGGAAACCGCTTCGAGCTGTGTGGCCGCTGATGCGGATTCGCTCTCGCA 71  
DB 345 CGGCTTCCTCGGTGCGGCGGCTCGGCTGTGCGGCTGTGAGAGACGAGCGCAATCGCGA 404  
QY 72 CAGCCCGGCGCCCACTGTGGAGNAGTGTGACGCTGTGAC 112  
DB 405 CATCTCGCGACACTGTGCGGCGGAGCGGCTCGCGCGCGAC 445

RESULT 14  
US-10-156-761-1/c  
; Sequence 1, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 1  
LENGTH: 9025608  
TYPE: DNA  
ORGANISM: Streptomyces avermitilis  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (4187715)  
OTHER INFORMATION: a, t, c, g, other or unknown  
US-10-156-761-1

Query Match 27.7% Score 36; DB 15; Length 9025608;  
Best Local Similarity 59.4%; Pred. No. 0.066;  
Matches 60; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 12 CGGCTACCCCGGAAACCGCTTGAGCTGTGGCGCTGATGGCGGATTCGCTCTCCGA 71  
Db 1994191 CGGCTTCGCTGGGCGGGGTCGGCTCGTGGCGCTGGAGGAGCGCGAAGCTGGCGCA 1994132  
Qy 72 CAGCCCCGGCCCCACCTGGGAGAGNAGTGTGACGCTCGTAC 112  
Db 1994131 CATCTCGCGCACTCTGCGGCGAGCGGCTCGGCCGAC 1994091

RESULT 15  
US-10-260-238-1034  
Sequence 1034, Application US/10260238  
Publication No. US20040016025A1  
GENERAL INFORMATION:  
APPLICANT: Budworth, Paul R.  
APPLICANT: Moughamer, Todd G.  
APPLICANT: Briggs, Steven P.  
APPLICANT: Cooper, Bret  
APPLICANT: Glazebrook, Jane  
APPLICANT: Goff, Stephen A.  
APPLICANT: Katagiri, Fumiyaki  
APPLICANT: Kreps, Joel  
APPLICANT: Provart, Nicholas  
APPLICANT: Rieke, Darrell  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
FILE REFERENCE: 60111-NP  
CURRENT APPLICATION NUMBER: US/10/260,238  
CURRENT FILING DATE: 2002-09-26  
PRIOR APPLICATION NUMBER: US 60/325,448  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/325,277  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/370,620  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 6077  
SEQ ID NO 1034  
LENGTH: 2412  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-10-260-238-1034

Query Match 27.2% Score 35.4; DB 17; Length 2412;

Best Local Similarity 60.6%; Pred. No. 0.2;  
Matches 57; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 3 CGCTACCTGGGCTACCCCGGGAACCGCTTGAGCTGTGGCGCTGATGGGGAATTCGT 62  
Db 1530 CCGCTCGCGGCGAGACGCGCAACATCTTGGCGCGGCTTGAAGATGGCGACCGGT 1589  
Qy 63 GCTCTCCGACAGCCCGGCGCCACCTGGAGNAG 96  
Db 1590 GGTACCTGAGCCCGGCTACCTGTGGAGCTG 1623

Search completed: June 7, 2005, 18:04:27  
Job time : 133.203 secs



and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."

## ORIGIN

Query Match 99.2%; Score 129; DB 1; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-23;  
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGCTACCTGCTACCCCGGGAACCGCTTCGAGCTGTGGCCCTGATGCGGATTCC 60  
 Db 1 TCCGCTACCTGCTACCCCGGGAACCGCTTCGAGCTGTGGCCCTGATGCGGATTCC 60

QY 61 GTGCTCTCGAGAGCCCGGCGCCCACTTGAGAGNAGTGATGACGCTTCGTGCGAG 120  
 Db 61 GTGCTCTCGAGAGCCCGGCGCCCACTTGAGAGNAGTGATGACGCTTCGTGCGAG 120

QY 121 GGACGCTGCT 130  
 Db 121 GGACGCTGCT 130

RESULT 2  
 AA005293 206 bp mRNA linear EST 07-MAY-1997  
 LOCUS zh93a11.r1 Soares fetal liver spleen INFIS\_S1 Homo sapiens cDNA  
 DEFINITION clone IMAGE:428828 5', mRNA sequence.  
 ACCESSION AA005293  
 VERSION AA005293.1 GI:1448755  
 KEYWORDS EST.  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
 1 (bases 1 to 206)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,  
 Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,  
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
 Maris, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
 Rolifing, T., Scheilenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
 Trevaeths, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,  
 and Warr, M.  
 Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)  
 97044478  
 8889549

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 849 Std Error: 0.00  
 Seq primer: mob.REGA+ET  
 High quality sequence stop: 126.

## FEATURES

source  
 1..206  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="CDB:1328597"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:428828"  
 /sex="male"  
 /dev\_stage="20 week-post conception fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares fetal liver spleen INFIS\_S1"  
 /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)  
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
 This is a subtracted version of the original Soares fetal  
 liver spleen INFIS library. 1st strand cDNA was primed  
 with a Pac I - oligo(dT) primer [5'

## ORIGIN

Query Match 57.2%; Score 74.4; DB 1; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 5.4e-09;  
 Matches 86; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 43 GCGGTAGTGGCGGATTCGCTCTCGACAGACCCCGCCCACTGAGG-AGNAGTGTG 101  
 Db 1 GCGGTAGTGGCGGATTCGCTCTCGACAGACCCCGCCCACTGAGGATGAGTGTG 60

QY 102 ACGCTGTGACCTTGCGAGGAGCGTGT 130  
 Db 61 ACGCTGTGACCTTGCGAGGAGCGTGT 89

RESULT 3  
 A1604990 339 bp mRNA linear EST 21-APR-1999  
 LOCUS v122b11.x1 Knowles Solter mouse unfertilized egg Mus musculus cDNA  
 DEFINITION clone IMAGE:836445 3', mRNA sequence.  
 ACCESSION A1604990  
 VERSION A1604990.1 GI:4614157  
 KEYWORDS EST.  
 ORGANISM Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 (bases 1 to 339)  
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,  
 Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R.,  
 Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
 Waterston, R. and Wilson, R.  
 The WashU-NCI Mouse EST Project 1999  
 Unpublished (1999)

TITLE  
 JOURNAL  
 COMMENT  
 Contact: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:496661  
 This clone was previously sequenced on the 5' end only, this new  
 data is from the 3' end  
 Seq primer: primer name ambiguous  
 High quality sequence stop: 319.

## FEATURES

source  
 1..339  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J x DBA/2J F1"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:836445"  
 /tissue\_type="unfertilized egg"  
 /lab\_host="DH10B"  
 /clone\_lib="Knowles Solter mouse unfertilized egg"  
 /note="Organ: unfertilized egg; Vector: pBluescribe  
 (modified); Site 1: MluI; Site 2: SalI; Cloned  
 unidirectionally from mRNA prepared from 5000 unfertilized  
 eggs. Primer: SalI (dT):  
 5'-CGGTGACCGTCGACGCTTTTCTTTT-3'. cDNAs were  
 cloned into the MluI/SalI sites of a modified pBluescribe  
 vector using commercial linkers (NEB). Average insert  
 size: 1.0 Kb."

## ORIGIN



Query Match 39.2%; Score 51; DB 1; Length 339;  
Best Local Similarity 67.2%; Pred. No. 0.0054;  
Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 1 TCCGCTACCTGCTACCCCGGGAACCGCTTCGAGCTGTGCGCGCTGATGCGGATTC 60  
DB 54 TCTCTCTCTGCGAAGCCGGGGCAATGCCCTGAGCTGTGTAACAGATGCGAGATAG 113  
QY 61 GTGCTCTCCGACAGACCCCGCCCACTTG-GAGNAGTGCTGACGCTGTAACCTTCGA 119  
DB 114 TTGCTCTCCAAAGACCAAGACTTCAGCTGAGCCCACTGTGATGCTCTCGGCTTCGCG 173  
QY 120 GGGAGCGCT 127  
DB 174 GGGAGCGCT 181

RESULT 4  
CA560084 448 bp mRNA linear EST 19-NOV-2002  
LOCUS K0267E03-5N NIA Mouse Unfertilized Egg cDNA Library (long) Mus  
DEFINITION musculus cDNA clone NIA:K0267E03 IMAGE:30051122 5', mRNA sequence.  
ACCESSION CA560084  
VERSION CA560084.1 GI:25104716  
KEYWORDS EST  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Piao Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Stagg, C.A.,  
Martin, P., Alba, K., Tanaka, T. and Ko, M.S.H.  
TITLE Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library  
(Long)

JOURNAL  
COMMENT Unpublished (2001)  
Other ESTs: K0267E03-3  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
Plate: K0267 row: B column: 03  
Seq primer: M13 Reverse  
High quality sequence stop: 448  
POLYA=No.

FEATURES  
source location/Qualifiers

1..448  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="tae:K0267E03-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0267E03 IMAGE:30051122"  
/tissue\_type="Unfertilized Egg"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse Unfertilized Egg cDNA Library  
(Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were  
extracted from a pool of 1488 unfertilized eggs.  
Double-stranded cDNAs were synthesized with an Oligo(dT)  
primer (Invitrogen):  
5'-pGACGATCTCTGATCGGAGCGCGCCCTTTTCTTTT-3',  
treated with T4 DNA polymerase, and purified by  
ethanol-precipitation. The cDNAs were ligated to  
lone-linker L1-Sal4, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using

Ex Tag polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.5 kb. The library was constructed  
by Yulan Piao (NIA)."

Query Match 39.2%; Score 51; DB 6; Length 448;  
Best Local Similarity 67.2%; Pred. No. 0.0054;  
Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 1 TCCGCTACCTGCTACCCCGGGAACCGCTTCGAGCTGTGCGCGCTGATGCGGATTC 60  
DB 133 TCTCTCTCTGCGAAGCCGGGGCAATGCCCTGAGCTGTGTAACAGATGCGAGATAG 192  
QY 61 GTGCTCTCCGACAGACCCCGCCCACTTG-GAGNAGTGCTGACGCTGTAACCTTCGA 119  
DB 193 TTGCTCTCCAAAGACCAAGACTTCAGCTGAGCCCACTGTGATGCTCTCGGCTTCGCG 252  
QY 120 GGGAGCGCT 127  
DB 253 GGGAGCGCT 260

RESULT 5  
CA561864 478 bp mRNA linear EST 19-NOV-2002  
LOCUS K0293B04-5N NIA Mouse Unfertilized Egg cDNA Library (long) Mus  
DEFINITION musculus cDNA clone NIA:K0293B04 IMAGE:3005583 5', mRNA sequence.  
ACCESSION CA561864  
VERSION CA561864.1 GI:25106519  
KEYWORDS EST  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Piao Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Stagg, C.A.,  
Martin, P., Alba, K., Tanaka, T. and Ko, M.S.H.  
TITLE Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library  
(Long)

JOURNAL  
COMMENT Unpublished (2001)  
Other ESTs: K0293B04-3  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
Plate: K0293 row: B column: 04  
Seq primer: M13 Reverse  
High quality sequence stop: 478  
POLYA=No.

FEATURES  
source location/Qualifiers

1..478  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="tae:K0293B04-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0293B04 IMAGE:3005583"  
/tissue\_type="Unfertilized Egg"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse Unfertilized Egg cDNA Library  
(Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were

extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo (dT) primer [Invitrogen: 5'-GGACTAGTCTAGTACGCGACGCGCCGCCCCCTTTTCTTTTCTTTT-3'], treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to *lon*-linker *l*L-Sal-I, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal-I-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with Sal I and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10 E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA). "

ORIGIN

Query Match	39.2%;	Score 51;	DB 6;	Length 478;
-------------	--------	-----------	-------	-------------

Best Local Similarity 67.23; Pred. NO. 0.0054;  
Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

Oy	1	TCGCGCACTCGGCTACCCCGGAAACGCTTCAGAGCTGGTGGCGCTATAGCGGAAATTC	60
Db	162	TCCCTCCTTCTCGCAAAAGCCGGGGCAATGCGCTCGAGAGCTGGTGAACAACATATGCGAATAG	221
Oy			
Oy	61	GTCCTCTCCGACACAGCCCGGCGCCACACTGG -GAGAGATGTGTGACGCTGTGTAACTTGGCA	119
Db	222	TTCGCTCTCCAAAGACCAAGACTTCAGCTGAAGCAACAATGTGATCTCTGAGCTTCGCG	281
Oy			
Oy	120	GGGAGCGT	127
Db	282	GGGAGCGT	289

RESULT 6	AI505208	LOCUS	DEFINITION
	AI505208	486 bp	mRNA
	vgj3g11.x1	Knowles Solter mouse 2 cell Mus musculus cDNA clone	linear
	IMAGE:1108004.3	mRNA sequence.	EST 11-MAR-1991

```

/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:1108004"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/clone_idb="Knowles Solter mouse 2 cell"
/note="Organ: embryo; Vector: pBluescribe (modified) from
M13: Mu1; Site 2: SalI; cloned unidirectionally from
mRNA prepared from 13,500 2-cell stage embryos. Primer:
SalI(dn): 5'-CGGTCAGCCGACCGATTTTTTTTTTTT-3'.  

CDNAS
were cloned into the Mu1/SalI sites of a modified
pBluescribe vector using commercial linkers (NEB).
Average insert size: 1.2 kb "
```

**ORIGIN**

Query Match	39.2%;	Score 51;	DB 1;	Length 486;
-------------	--------	-----------	-------	-------------

Best Local Similarity 87.2%; Prec.No. 0.0034;  
Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1

Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1.

Oy	1	TCGCGCTACCTGGGCTACCCCGGAAACGGCTTCGAGCTGGGCGCGTATGCGGATTC	60
		283	TCCTCTTTCGGAAAGCCGGGGCAATGCCCTGGAGCTGGTGAACAAGATGGCAGATTAAG
Oy	61	GTGCTTCCGACAGCCCGGCCCCCACTGG-GAGNAGTGTGAGCGTCGAGACCTTGCA	119
Db	343	TTGCTCTCCAAAGACCAGATTCACGCTGAGCCAACTGGTAGTCTCTCGGCTTCGCG	402
Oy	120	GGGAGCCT	127
Db	403	GGGAGCCT	410

RESULT	7
CAS60564	
LOCUS	524 bp mRNA linear EST 19-NOV-2001
DEFINITION	CAS60564 CA56056-5N NIA Mouse Unfiltered Egg cDNA Library (long) Mus
	K0274C06-5N NIA:K0274C06 IMAGE:30051773 5', mRNA sequence.

**FEATURES**  
**SOURCE**

FEATURES	
source	location/Qualifiers
1..486	
/organism="Mus musculus"	/tissue_type="Unfertilized Egg"
/mol_type="mRNA"	/lab_host="DH10B"
	/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://1gsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]]. Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo (dt) primer [Invitrogen]:  
5'-GCACGATGTTCTGATCGGAGCGCGCCCTTTT-3',  
treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

Query Match 39.2%; Score 51; DB 6; Length 524;

Best Local Similarity 67.2%; Pred. No. 0.0054; Mismatches 41; Indels 1; Gaps 1;

Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

Qy 1 TCCGCTTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGGCCGTGATGCCGATTC 60  
Db 127 TCCCTCTTCTGCGAAGCGGGGCAATGCGCTGAGCTGTGTAACAGATGCGACATAG 186  
Qy 61 GTGCTCTCCGACAGCCCGCCGACCTTG-GAGNAGTGTGACGCTGTCGACTTCGCA 119  
Db 187 TTGCTCTCCAAAGCAAGACTTCACTGAGACCAACTGTGATGCTCTGCGCTTCGCG 246  
Qy 120 GGGACGCT 127  
Db 247 GGGACGCT 254

RESULT 8  
CA557550 588 bp mRNA linear EST 19-NOV-2002  
LOCUS K0229B10-5N NIA Mouse Unfertilized Egg cDNA Library (Long) Mus  
DEFINITION Musculus cDNA clone NIA:K0229B10 IMAGE:30047445 5', mRNA sequence.  
ACCESSION CA557550  
VERSION CA557550.1 GI:25102015  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 588)  
AUTHORS Mamalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Piao Y., Kargul G.J., Dudekula D.B., Qian Y., Luo A., Steggy C.A., Martin, P., Alida, K., Tanaka, T. and Ko, M.S.H.  
TITLE Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library (Long)  
JOURNAL Unpublished (2001)  
COMMENT Other ESTs: K0229B10-3  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@1gsun.grc.nia.nih.gov  
Plate: K0229 row: B column: 10  
Seg primer: M13 Reverse  
High quality sequence stop: 588  
POLYA=No.

FEATURES  
SOURCE 1..588  
Location/Qualifiers

/organism="Mus musculus"

/mol type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="taxon:10090"  
/db\_xref="taxon:10090"  
/clone="NIA:K0229B10 IMAGE:30047445"  
/tissue\_type="Unfertilized Egg"  
/lab host="DH10B"  
/clone\_lib="NIA Mouse Unfertilized Egg cDNA Library (Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://1gsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]]. Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo (dt) primer [Invitrogen]:  
5'-GCACGATGTTCTGATCGGAGCGCGCCCTTTT-3',  
treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

Query Match 39.2%; Score 51; DB 6; Length 588;

Best Local Similarity 67.2%; Pred. No. 0.0054; Mismatches 41; Indels 1; Gaps 1;

Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

Qy 1 TCCGCTTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGTGGCCGTGATGCCGATTC 60  
Db 172 TCCCTCTTCTGCGAAGCGGGGCAATGCGCTGAGCTGTGTAACAGATGCGACATAG 231  
Qy 61 GTGCTCTCCGACAGCCCGCCGACCTTG-GAGNAGTGTGACGCTGTCGACTTCGCA 119  
Db 232 TTGCTCTCCAAAGCAAGACTTCACTGAGACCAACTGTGATGCTCTGCGCTTCGCG 291  
Qy 120 GGGACGCT 127  
Db 292 GGGACGCT 299

RESULT 9  
BG080862 626 bp mRNA linear EST 17-DEC-2003  
LOCUS H3058C03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
DEFINITION H3058C03 5', mRNA sequence.  
ACCESSION BG080862  
VERSION BG080862.2 GI:40014930  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 626)  
AUTHORS Mamalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grubovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagajaya, R., Dol, H., Wood, W.H., III, Becker, K.G. and Ko, M.S.H.

TITLE Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)  
MEDLINE 20381348  
PUBMED 10922068  
COMMENT On Jan 26, 2001 this sequence version replaced gi:12563419.

Other ESTs: H3058C03-3  
 Contact: George J. Kargul  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 This clone set has been freely distributed to the community. Please  
 visit <http://1gsun.grc.nia.nih.gov/cDNA/15k.html> for details.  
 Plate: H3058 row: C column: 03  
 Seq primer: -21M13 Reverse  
 High quality sequence stop: 626  
 POLYA=No.

# FEATURES

source

Location/Qualifiers

1..626

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taeST:H3058C03-5"

/db\_xref="taxon:10090"

/clone="H3058C03"

/sex="Clones arrayed from a variety of cDNA libraries"

/dev\_stage="Clones arrayed from a variety of cDNA libraries"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse 15K cDNA Clone Set"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-NotI primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

## ORIGIN

Query Match 39.2%; Score 51; DB 4; Length 626;  
 Best Local Similarity 67.2%; Pred. No. 0.0054;

Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

```

1 TCCGCTTACCTGCGTACCCCGGGAACCGCTTCGAGCTGTGGCGCGGATTC 60
|||||
174 TCCCTCTTCTGGAAGCCGGGCAATCGCTGAGCTGTGAACAGATGGCAGATTA 233
61 GTGCTCTCCGAGAGCCCGGCGCCCACTGG-GAGNAGTGATGACGCTCGTACCTTCGA 119
|||||
234 TTGCTCTCCAAAGACCAAGACTTCGAGCTGTGAGCAACTGTATGTCTTGCGCTTCGCG 293
120 GGGACGCT 127
|||||
294 GGGACGCT 301

```

## RESULT 10

CF915355 649 bp mRNA linear EST 05-NOV-2003  
 LOCUS B0978E08-5 NIA Mouse Unfertilized Egg cDNA library (Long 1) Mus  
 DEFINITION muscular cDNA clone NIA:B0978E08 IMAGE:30479767 5', mRNA sequence.  
 ACCESSION CF915355  
 VERSION CF915355.1 GI:38186557  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

# REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 649)  
 Piao, Y., Ko, N.T., Lam, M.K. and Ko, M.S.H.  
 Construction of long-transcript enriched cDNA libraries from  
 submicrogram amounts of total RNAs by a universal PCR amplification  
 method  
 Genome Res. 11 (9), 1553-1558 (2001)  
 21429098  
 11544199  
 Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: B0978 row: E column: 08  
 Seq primer: M13 Reverse  
 High quality sequence stop: 649  
 POLYA=No.

# FEATURES

source

Location/Qualifiers

1..649

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taeST:B0978E08-5"

/db\_xref="taxon:10090"

/clone="NIA:B0978E08 IMAGE:30479767"

/dev\_stage="Unfertilized Eggs"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse Unfertilized Egg cDNA Library (Long 1)"

/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://1gsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). (PMID: 11544199). Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen): 5'-TGACTGATTCAGATCGGAGGCGCCCTTTT-3', treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

## ORIGIN

Query Match 39.2%; Score 51; DB 7; Length 649;  
 Best Local Similarity 67.2%; Pred. No. 0.0054;

Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

```

1 TCCGCTTACCTGCGTACCCCGGGAACCGCTTCGAGCTGTGGCGCGGATTC 60
|||||
174 TCCCTCTTCTGGAAGCCGGGCAATCGCTGAGCTGTGAACAGATGGCAGATTA 233
61 GTGCTCTCCGAGAGCCCGGCGCCCACTGG-GAGNAGTGATGACGCTCGTACCTTCGA 119
|||||
234 TTGCTCTCCAAAGACCAAGACTTCGAGCTGTGAGCAACTGTATGTCTTGCGCTTCGCG 293
120 GGGACGCT 127
|||||
294 GGGACGCT 301

```

RESULT 11  
BY735639  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BY735639 736 bp mRNA linear EST 17-DEC-2002  
BY735639 RIKEN full-length enriched, 1 cell embryo Mus musculus  
cDNA clone 10C0031F16 5', mRNA sequence.  
BY735639.1 GI:27148766  
EST.  
Mus musculus  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 736)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bul, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
Chochoa, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,  
Flachner, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Glasi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, J.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurachkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Meglert, D.R.,  
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Nunata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G.,  
Petrovsky, N., Pilla, R., Pontius, J.U., Qi, D., Ramchandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Sempke, C.A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wyszynski, B., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carrinci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, U., Aizawa, K.,  
Aizawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
Shinawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
22354683  
12466851  
PUBMED  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Saitoh-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Adachi, J., Aizawa, K., Akimura, T., Aizawa, T., Carrinci, P.,  
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,  
Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,  
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,  
Nemura, K., Numazaki, R., Ohno, M., Otsu, N., Saito, R., Sakazume, N.,  
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,  
Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

FEATURES  
source  
Location/Qualifiers  
1..736  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="taxon:10090"  
/clone="10C0031F16"  
/cell\_type="1 cell"  
/dev\_stage="1 cell embryo"  
/clone\_11b="RIKEN full-length enriched, 1 cell embryo"

ORIGIN  
Query Match 39.2% Score 51; DB 6; Length 736;  
Best Local Similarity 67.2% Pred. No. 0.0053;  
Matches 86; Conservative 0; Mismatches 41; Indels 1; Gaps 1;  
1 TCCGCTTACCTCGGCTACCCGGGAGACCGCTTCCAGCTGTGGCGCTGATGCGGATTC 60  
185 TCTCTCTTCTCGGAAGCGCGGCGCATGCTCGAGCTGGTGGAAACAGATGACAGTAAG 244  
61 GTGCTCTCGGACAGCCCGGCGCCACCTGG-GAGNAGTGTGAGCGCTGTCCTTGGCA 119  
245 TTGCTCTCAAGACCAAGACTTCACTGAGACCACTGTGATCTCTGCTTCGCG 304  
120 GGGAGCGCT 127  
305 GGGAGCGCT 312

RESULT 12  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CO799635 758 bp mRNA linear EST 05-AUG-2004  
AGNCOURT 30246635 NIH MGC 256 Mus musculus cDNA clone  
IMAGE:30937853 5', mRNA sequence.  
CO799635.1 GI:50987815  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 758)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: dsgerb@remail.nih.gov  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
cDNA Library Preparation: Express Genomics  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNLT at:  
<http://image.llnl.gov>  
Plate: NDAMI195 row: n column: 06  
High quality sequence stop: 623.  
Location/Qualifiers  
1..758  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30937853"  
/lab\_host="DH10B TONa"  
/clone\_11b="NIH\_MGC\_256"

/note="Organ: ovocyte; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; cDNA was primed using oligo-dT primer: 5'-GACAGTTGTCATGCGAGCGCGGCC (T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb resulted in an average insert size of 1.2 kb. This is a primary library (normalized primary library is NIH MGC 257) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH MGC library"

Query Match	39.2%;	Score 51;	DB 7;	Length 758;
Best Local Similarity	67.2%;	Pred. No. 0.0053;		
Matches	86;	Conservative	0;	Mismatches 41; Indels 1; Gaps 1

Qy	1	TCGCGCTACCTCGGCTACCCCGGGAACCGCTTCGACCTGGTGGCGCTGATATGGCGGATATCC	60
Db	207	TCCTCTCTTCTGTGGAAAGCCGGGGCAATCCCTCGAGCTGTGTAAACAGATGGAGATTAAG	266
Qy	61	GTGCTCTCGACAGCCGCCGAGCCCACTTG3-GAGNAGTGGTGA CGCTCGTGA CTTTGCA	119
Db	267	TTGCTCTCCAAAGACCAAGACTTTCAGCTGGAGCCAACTGGTATGCTCCTGGGCTTTCGCG	326
Qy	120	GGGACGCT	127
Db	327	GGGACGCT	334

RESULT	13
CO808612	
LOCUS	763 bp mRNA linear EST 06-AUG-2004
DEFINITION	AGENCOURT 30256728 NIH_MGC_256 mus musculus cDNA clone IMAGE:30938958 5', mRNA sequence.
ACCESSION	CO808612
VERSION	CO808612.1 GI:51025466

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
AUTHORS	1 (bases 1 to 763)
TITLE	Nih-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999)
	Contact: Daniela S. Gerhard, Ph.D.

Bldg. 31 Rml0A07 Bethesda, MD 20892  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Kathleen Horner, Stanford University  
 CDNA Library Preparation: Express Genomics  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: NDAM1198 row: 1 column: 07  
 High quality sequence stop: 540.

```

FEATURES
source
location/Qualifiers
1..763
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30938958"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_256"
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV;
Site 2: NotI; cDNA was primed using oligo-dT primer:
5'-pGACTGATCTCTGATCGAGCGGCGGCC(7)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb
resulted in an average insert size of 1.2 kb. This is a
primary library (normalized primary library is NIH MGC 2577)
and was constructed by Express Genomics (Frederick, MD).
Note: this is a NIH_MGC library"

```

ORIGIN			
Query Match	39.2%	Score 51;	DB 7;
Best Local Similarity	67.2%	Pred. No. 0.0053;	Length 763;
Matches	86;	Mismatches	41;
		Indels	1;
		Gaps	1;

OY 1 TCBCGCTACCTCGGCTACCCCGGGAGCCGTTTCAGCTGTGGCGCATGGCGATTC 60  
| | | | |  
Db 72 TCCCTCTTCTCGGAAGGCCGGGCAGATCGCTGAGCGTGGAAGAAGATGGCAGATAAG 131  
| | | | |  
OY 61 GNGCTCTTCGAGCAGCCCCCGGCCCACTGG- GAGNAGTGTGACGCTCGTAGCTTGCGCA 119  
| | | | |  
Db 132 TTGGCTCTCGAAGAACAAGACTTCAAGTTGGAGGCCAACTGGTGAATGCTCTTGCGCTGCGC 191  
| | | | |  
OY 120 GGAGCGCT 127  
| | | | |  
Db 192 GGGAGCGT 199

RESULT 14					
LOCUS	COT97157	770 bp	mRNA	linear	EST 05-AUG-2004
DEFINITION	AGNCOURT.30256532 NIH_MGC_256 Mus musculus cDNA clone IMAGE:30936954 5', mRNA sequence.				
ACCESSION	COT97157				
VERSION	COT97157.1	GI:50985337			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1  
(bases 1 to 770)  
NIH-MGC <http://mgc.ncbi.nlm.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Biotechnology Acquisition

National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cga@bbs-rcmail.nih.gov  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: WGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAMI193 row: h column: 19  
High quality sequence start: 23  
High quality sequence stop: 417.

```

source
1. .7/0
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30936954"
/lab_host="DH10B T0md"
/clone_id="NIH_MGC_256"
/notice="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV;
Site 2: NotI; cDNA was primed using oligo-dT primer:
5'-pGACGACTTCGACATCCGACGAGCGCGCCGCTT25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb
resulted in an average insert size of 1.2 kb. This is a
primary library (normalized primary library is NIH_MGC_257)
and was constructed by Express Genomics (Frederick, MD).
Note: this is a NIH_MGC library"

```

Query Match	39.2%	Score 51;	DB 7;	Length 770;
Best Local Similarity	67.2%;	Pred. No. 0.0053;		
Matches	86;	Conservative	0;	Mismatches 41; Indels 1; Gaps 1.

QY 1 TCCGCTACTCGCTACCCGGAGACCGCTTCAGCTGTGCGCGCTGATGAGCGATTCC 60

```

Db      323 TCCTCTTCTGCGAAGCCGGGCAATCGCTGAGCTGGTGAACGATGGCAGATTAG 382
Qy      61 GTGCTCTCGACAGACCCCGCCGCTGAG-AGNAGTGTGACGCTCGGACCTTCGCA 119
Db      383 TTGCTCTCCAAAGACCAGACTTCAGCTGAGCCCACTGTGTGCTCTCGGCTTCGCG 442
Qy      120 GGGACGCT 127
Db      443 GGGACGCT 450

```

```

Qy      120 GGGACGCT 127
Db      318 GGGACGCT 325

```

Search completed: June 7, 2005, 17:48:14  
Job time : 794.354 secs

## RESULT 15

CO814648

LOCUS

DEFINITION CO814648 772 bp mRNA linear EST 06-AUG-2004

AGENCOURT\_30246837 NIH\_MGC\_256 Mus musculus cDNA clone

IMAGE:30937746 5', mRNA sequence.

CO814648

VERSION CO814648.1 GI:51033274

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1. (bases 1 to 772)

NIH-MGC http://mgs.nci.nih.gov/

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgeabs-remail.nih.gov

Tissue Procurement: Dr. Kathleen Horner, Stanford University

cDNA library Preparation: Express Genomics

cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM1195 row: 1 column: 19

High quality sequence start: 5

High quality sequence stop: 666.

Location/Qualifiers

FEATURES

source

1..772

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:30937746"

/lab\_host="DH10B Tona"

/clone\_lib="NIH\_MGC\_256"

/note="Organ: oocyte; Vector: pExpress-1; Site\_1: EcoRV;

Site\_2: NotI; cDNA was primed using oligo-dT primer:

5'-pGACTGTTCTAGATCGAGCGGCCGCTT25-3' and cloned into

the EcoRV/NotI sites of pExpress-1. Size-selection &gt;0.5 kb

resulted in an average insert size of 1.2 kb. This is a

primary library (normalized primary library is NIH\_MGC\_257)

and was constructed by Express Genomics (Frederick, MD).

Note: this is a NIH\_MGC library"

## ORIGIN

Query Match 39.2%; Score 51; DB 7; Length 772;

Best Local Similarity 67.2%; Pred. No. 0.0053; Mismatches 41; Indels 1; Gaps 1;

Matches 86; Conservative 0;

Qy 1 TCGCCTTACCTCGGCTACCCCGGGAACCCCTTGAAGCTGAGCGGCGGATGCC 60

Db 198 TCCTCTTCTGCGAAGCCCGGCGCAATCGCTGAGCTGTGAACAGATGGCAGATTAG 257

Qy 61 GTGCTCTCGACAGACCCCGCCGCTGAG-AGNAGTGTGACGCTCGGACCTTCGCA 119

Db 258 TTGCTCTCCAAAGACCAGACTTCAGCTGAGCCCACTGTGTGCTCTCGGCTTCGCG 317

**This Page Blank (uspto)**